

92394

STIC-Biotech/ChemLib

From: Spector, Lorraine  
Sent: Friday, April 25, 2003 8:06 AM  
To: STIC-Biotech/ChemLib  
Subject: SEARCH request for Serial No. 09/674377

STIC,

Please search SEQ ID NO:1

- pending
- issued
- commercial

Thanks.

Lorraine Spector  
703-308-1793  
U.S. Patent and Trademark Office  
Art Unit 1647  
lorraine.spector@uspto.gov  
CM1-10B11  
Mailbox 10-B19

Edward Hart  
Technical Info. Specialist  
STIC/Biotech  
CMI 6B02 Tel: 305-9203

Searcher: \_\_\_\_\_  
Phone: \_\_\_\_\_  
Location: \_\_\_\_\_  
Date Picked Up: 4/25/03  
Date Completed: 4/26/03  
Searcher Prep/Review: \_\_\_\_\_  
Clerical: \_\_\_\_\_  
Online time: \_\_\_\_\_

TYPE OF SEARCH:  
NA Sequences: \_\_\_\_\_  
AA Sequences: 1  
Structures: \_\_\_\_\_  
Bibliographic: \_\_\_\_\_  
Litigation: \_\_\_\_\_  
Full text: \_\_\_\_\_  
Patent Family: \_\_\_\_\_  
Other: \_\_\_\_\_

VENDOR/COST (where applic.)  
STN: \_\_\_\_\_  
DIALOG: \_\_\_\_\_  
Questel/Orbit: \_\_\_\_\_  
DRLink: \_\_\_\_\_  
Lexis/Nexis: \_\_\_\_\_  
Sequence Sys.: 72  
WWW/Internet: \_\_\_\_\_  
Other (specify): \_\_\_\_\_

**THIS PAGE BLANK (USPTO)**

## Spector, Lorraine

---

**From:** Spector, Lorraine  
**Sent:** Friday, April 25, 2003 8:06 AM  
**To:** STIC-Biotech/ChemLib  
**Subject:** SEARCH request for Serial No. 09/674377

STIC,

Please search SEQ ID NO:1

-pending

-issued

-commercial

Thanks.

Lorraine Spector  
703-308-1793  
U.S. Patent and Trademark Office  
Art Unit 1647  
lorraine.spector@uspto.gov  
CM1-10B11  
Mailbox 10-B19

**THIS PAGE BLANK (USPTO)**

# BioTech-Chem Library

## Search Results

### Feedback Form (Optional)



Scientific & Technical Information

The search results generated for your recent request are attached. If you have any questions or comments (compliments or complaints) about the scope or the results of the search, please contact *the BioTech searcher* who conducted the search *or contact*:

Mary Hale, Supervisor, 30  
CM-1 Room 1E01

---

#### *Voluntary Results Feedback Form*

➤ *I am an examiner in Workgroup: 1647 (Example: 1610)*

➤ *Relevant prior art found, search results used as follows:*

- ☐ 102 rejection
- ☐ 103 rejection
- ☐ Cited as being of interest.
- ☐ Helped examiner better understand the invention.
- ☐ Helped examiner better understand the state of the art in their technology.

*Types of relevant prior art found:*

- ☐ Foreign Patent(s)
- ☐ Non-Patent Literature  
(journal articles, conference proceedings, new product announcements etc.)

➤ *Relevant prior art not found:*

- ☐ Results verified the lack of relevant prior art (helped determine patentability).
- ☐ Search results were not useful in determining patentability or understanding the invention.

**Other Comments:**

---

Drop off completed forms at the Circulation Desk CM-1, or send to Mary Hale, CM1-1E01 or [mary.hale@uspto.gov](mailto:mary.hale@uspto.gov)

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 25, 2003, 15:44:58 ; Search time 77 Seconds  
(without alignments)  
773.546 Million cell updates/sec

Title: US-09-674-377B-1

Perfect score: 2604

Sequence: 1 ERKRRNTIHEFKSAKTLI.....IPWDYCPISRCGGTPTIV 447

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_101002.\*

- 1: /SID22/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.\*
- 2: /SID22/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.\*
- 3: /SID22/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.\*
- 4: /SID22/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.\*
- 5: /SID22/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.\*
- 6: /SID22/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.\*
- 7: /SID22/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.\*
- 8: /SID22/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.\*
- 9: /SID22/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.\*
- 10: /SID22/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.\*
- 11: /SID22/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.\*
- 12: /SID22/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.\*
- 13: /SID22/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.\*
- 14: /SID22/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.\*
- 15: /SID22/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.\*
- 16: /SID22/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.\*
- 17: /SID22/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.\*
- 18: /SID22/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.\*
- 19: /SID22/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.\*
- 20: /SID22/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.\*
- 21: /SID22/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.\*
- 22: /SID22/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.\*
- 23: /SID22/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2601	99.9	697	21	AA198485 Hepatocyte growth
2	2601	99.9	697	21	AA198485 Sequence of a pept
3	2601	99.9	697	22	AA004275 Nuclear ligand Pep
4	2601	99.9	697	22	AA045838 Nucleic acid trans
5	2601	99.9	728	13	AA020005 Human hepatocyte g
6	2601	99.9	728	13	AA025676 Recombinant human
7	2601	99.9	728	14	AA040862 Competitive inhibi
8	2601	99.9	728	14	AA040863 Competitive inhibi
9	2601	99.9	728	14	AA042062 Vascular endotheli
10	2601	99.9	728	17	AAW00338 Human hepatic pare

11	2601	99.9	728	17	AAW00340 Wild type hepatocy
12	2601	99.9	728	17	AA087522 Mutant hepatocyte
13	2601	99.9	728	17	AA087523 Mutant hepatocyte
14	2601	99.9	728	17	AA087524 Mutant hepatocyte
15	2601	99.9	728	17	AA087525 Mutant hepatocyte
16	2601	99.9	728	19	AAW59222 Human leukocyte-de
17	2601	99.9	728	19	AAW58696 Human hepatocyte g
18	2601	99.9	728	19	AAW42998 Recombinant human
19	2601	99.9	728	19	AAW39207 Human hepatocyte g
20	2599	99.8	447	21	AAW57173 N-terminal region
21	2595	99.7	727	12	AA010656 Hepatic parenchyma
22	2595	99.7	728	12	AA014243 Human hepatocyte g
23	2595	99.7	728	12	AA015623 Human leukocyte-de
24	2595	99.7	728	14	AA039521 Hepatocyte growth
25	2591	99.5	728	20	AAW85529 Human hepatocyte g
26	2588	99.4	728	20	AAW85532 Human hepatocyte g
27	2587	99.3	728	20	AAW85531 Human hepatocyte g
28	2586	99.3	728	20	AAW85530 Human hepatocyte g
29	2577	99.0	728	15	AA047227 Hepatocyte growth
30	2565	98.5	728	12	AA012792 Human hepatocyte g
31	2565	98.5	728	14	AA052940 Human hepatocyte g
32	2565	98.5	728	14	AA052941 Human hepatocyte g
33	2565	98.5	728	14	AA052942 Human hepatocyte g
34	2565	98.5	728	14	AA052943 Human hepatocyte g
35	2565	98.5	728	14	AA052944 Human hepatocyte g
36	2565	98.5	728	14	AA052945 Human hepatocyte g
37	2565	98.5	728	14	AA052946 Human hepatocyte g
38	2565	98.5	728	14	AA052947 Human hepatocyte g
39	2565	98.5	728	14	AA052948 Human hepatocyte g
40	2565	98.5	728	19	AAW48221 Human recombinant
41	2565	98.5	728	20	AAW97041 A human hepatocyte
42	2565	98.5	728	22	AA068090 Human hepatocyte g
43	2565	98.5	728	22	AA068090 Amino acid sequenc
44	2563.5	98.4	723	11	AA07144 Tumour cytotoxic f
45	2563.5	98.4	723	12	AA015624 Human leukocyte-de

ALIGNMENTS

RESULT 1

AA198485  
ID AA198485 standard; Protein; 697 AA.

XX AA198485;  
XX AA198485

DT 31-JUL-2000 (first entry)

XX Hepatocyte growth factor Pep 20 used in nucleic acid transporter system.

DE Transporter system; nucleic acid delivery; gene therapy; cancer;  
KW carcinogenesis; cardiovascular disease; infection.

XX Synthetic.

XX US6033884-A.

XX 07-MAR-2000.

XX 14-DEC-1993; 93US-0167641.

XX 20-MAR-1992; 92US-0855389.

XX 19-MAR-1993; 93WO-US02725.

XX (BAYU) BAYLOR COLLEGE MEDICINE.

XX Gottchalk S, Sparrow J, Cristiano RJ, Woo SLC, Smith LC;

XX WPI; 2000-281993/24.

XX System for transporting nucleic acid into cells, useful e.g. in gene  
therapy and for generating transgenic animals, comprises binding agent  
linked to nucleic acid, surface ligand and lytic agent

XX Disclosure; Figure 23A; 108pp; English.

PS The present invention relates to a transporter system for delivering

XX nucleic acid to a cell. The system comprises a nucleic acid binding

CC complex, consisting of a binding molecule bonded non-covalently to the

CC nucleic acid, and covalently to a surface ligand, and a lytic agent. The

CC binding molecule is spermine or a spermidine derivative. Nucleotide

CC sequences AAA36633-A36652 and peptide sequences AAY98456-Y98500 are used

CC in the construction of the transporter system of the invention. The

CC transporter system is used in gene therapy, particularly to deliver

CC nucleic acids to hepatocytes, muscle cells or bone forming cells, e.g

CC for treating cardiovascular disease, cancer, and infection. The

CC transporter systems are also used to create transgenic animals (as models

CC for human carcinogenesis or disease or for drug testing). Other uses

CC include transforming cells to produce proteins, or transfecting cells in

CC vitro to study the function of the nucleic acid. The use of a surface

CC ligand allows specific targeting of selected cells and tissues. The lytic

CC agent provides for release of the nucleic acid into the cellular

CC interior, from endosomes, without requiring endosomal or lysosomal

CC degradation.

XX Sequence 697 AA;

SQ

Query Match 99.9%; Score 2601; DB 21; Length 697;

Best Local Similarity 99.8%; Pred. No. 1.3e-165;

Matches 446; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ERKRRNTIHEFKSAKTLTIKIDPALKIKTKKVTADQCANRCTRNKGLPFTCKAFVFK 60

DB 1 QKRRNTIHEFKSAKTLTIKIDPALKIKTKKVTADQCANRCTRNKGLPFTCKAFVFK 60

QY 61 ARKQCLWFPNSMSGVKKKEFGHEFDLYENKDYIRNCIIIGKGRSYKGTYSITKSGIKCOP 120

DB 61 ARKQCLWFPNSMSGVKKKEFGHEFDLYENKDYIRNCIIIGKGRSYKGTYSITKSGIKCOP 120

QY 121 WSSMIPHEHSFLPSSYRGKDLQENYCRNPRGEGGWCFTSNPEVRYEVCIDIPOCSEVEC 180

DB 121 WSSMIPHEHSFLPSSYRGKDLQENYCRNPRGEGGWCFTSNPEVRYEVCIDIPOCSEVEC 180

QY 181 MTCNGESYRGKDLQENYCRNPRGEGGWCFTSNPEVRYEVCIDIPOCSEVEC 240

DB 181 MTCNGESYRGKDLQENYCRNPRGEGGWCFTSNPEVRYEVCIDIPOCSEVEC 240

QY 301 QYPHEHDMTPENFKCDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDMSHGQDC 360

DB 301 QYPHEHDMTPENFKCDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDMSHGQDC 360

QY 361 YRGNGKNTYGNLSQTRSLGTCSDMKDNLHRIHFWEPDASKLNEYCRNPDGDAHGPW 420

DB 361 YRGNGKNTYGNLSQTRSLGTCSDMKDNLHRIHFWEPDASKLNEYCRNPDGDAHGPW 420

QY 421 CYTGNPLIPWDYCPISRCGGTTPTIV 447

DB 421 CYTGNPLIPWDYCPISRCGGTTPTIV 447

RESULT 2

AAY59030

ID AAY59030 standard; peptide: 697 AA.

XX AAY59030;

AC AAY59030;

XX 07-MAR-2000 (first entry)

DT

XX Sequence of a peptide ligand Pep20.

DE

XX Nucleic acid transport system; NTS; cell surface receptor; cytolysis;

KW nuclear membrane; lysis moiety; transgenic animal; human disease;

KW

XX nucleic acid delivery; cancer.

OS Synthetic.

XX US5994109-A.

PN 30-NOV-1999.

XX 03-JUN-1995; 95US-0460890.

XX 14-DEC-1993; 93US-0167641.

XX 20-MAR-1992; 92US-0855389.

PR 19-MAR-1993; 93WO-US02725.

PR 14-DEC-1993; 93US-0167641.

XX (BAYU ) BAYLOR COLLEGE MEDICINE.

PA Woo SLC, Cristiano RJ, Gottchalk S, Sparrow J, Smith LC;

PI WPI; 2000-038262/03.

XX Nucleic acid transport system, useful for creating transgenic animals

DR for assessing human disease such as cancer in an animal model

XX Disclosure; Fig 23A; 107pp; English.

PS The invention relates to a nucleic acid transport system (NTS) for

XX delivering nucleic acid into a cell. The NTS contains but is not limited

CC to 5 components: (a) the nucleic acid or a macromolecule to be delivered;

CC (b) a moiety that recognizes and binds to a cell surface receptor or

CC antigen or is capable of entering a cell through cytolysis; (c) a nucleic

CC acid or macromolecular molecule binding moiety; (d) a moiety that is

CC capable of moving or initiating movement through a nuclear membrane; and/

CC or (e) a lysis moiety that enables the transport of the entire complex

CC from the cell surface directly into the cytoplasm of the cell. The NTS

CC delivers nucleic acid into the cellular interior as well as the nucleus

CC of specific cells. The NTS can be used to treat disorders by targeting

CC specific nucleic acid accordingly. The NTS can also be used to create

CC transgenic animals for assessing human disease, such as cancer, in an

CC animal model. The NTS can be used in vitro with tissue culture cells

CC which allows the role of various nucleic acids to be studied by targeting

CC specific expression into specifically targeted tissue culture cells. The

CC lysis agent within the NTS avoids the problem of endosomal/lysosomal

CC degradation.

XX Sequence 697 AA;

SQ

Query Match 99.9%; Score 2601; DB 21; Length 697;

Best Local Similarity 99.8%; Pred. No. 1.3e-165;

Matches 446; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ERKRRNTIHEFKSAKTLTIKIDPALKIKTKKVTADQCANRCTRNKGLPFTCKAFVFK 60

DB 1 QKRRNTIHEFKSAKTLTIKIDPALKIKTKKVTADQCANRCTRNKGLPFTCKAFVFK 60

QY 61 ARKQCLWFPNSMSGVKKKEFGHEFDLYENKDYIRNCIIIGKGRSYKGTYSITKSGIKCOP 120

DB 61 ARKQCLWFPNSMSGVKKKEFGHEFDLYENKDYIRNCIIIGKGRSYKGTYSITKSGIKCOP 120

QY 121 WSSMIPHEHSFLPSSYRGKDLQENYCRNPRGEGGWCFTSNPEVRYEVCIDIPOCSEVEC 180

DB 121 WSSMIPHEHSFLPSSYRGKDLQENYCRNPRGEGGWCFTSNPEVRYEVCIDIPOCSEVEC 180

QY 181 MTCNGESYRGKDLQENYCRNPRGEGGWCFTSNPEVRYEVCIDIPOCSEVEC 240

DB 181 MTCNGESYRGKDLQENYCRNPRGEGGWCFTSNPEVRYEVCIDIPOCSEVEC 240

QY 241 YTLDPHTRWEYCAIKTCADNTMNDTVPLETTCIOQGEYRGTVNTIWNIGIPCORWDS 300

DB 241 YTLDPHTRWEYCAIKTCADNTMNDTVPLETTCIOQGEYRGTVNTIWNIGIPCORWDS 300

QY 301 QYPHEHDMTPENFKCDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDMSHGQDC 360

DB 301 QYPHEHDMTPENFKCDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDMSHGQDC 360

QY 361 YRGNGKNTYGNLSQTRSLGTCSDMKDNLHRIHFWEPDASKLNEYCRNPDGDAHGPW 420

DB 361 YRGNGKNTYGNLSQTRSLGTCSDMKDNLHRIHFWEPDASKLNEYCRNPDGDAHGPW 420

QY 421 CYTGNPLIPWDYCPISRCGGTTPTIV 447

DB 421 CYTGNPLIPWDYCPISRCGGTTPTIV 447

RESULT 2

AAY59030

ID AAY59030 standard; peptide: 697 AA.

XX AAY59030;

AC AAY59030;

XX 07-MAR-2000 (first entry)

DT

XX Sequence of a peptide ligand Pep20.

DE

XX Nucleic acid transport system; NTS; cell surface receptor; cytolysis;

KW nuclear membrane; lysis moiety; transgenic animal; human disease;

KW



Db 301 QYPHEHDMTPENFKCKDLRENYCRNPDSGSPWCFTTDPNIRVGYCSQIPNCDSHGGDC 360  
QY 361 YRGNGKNTMGNSQTRSGLTCSMDKNMEDLHRHIFWEPDASKLNNENYCRNPDDDAHGPW 420  
Db 361 YRGNGKNTMGNSQTRSGLTCSMDKNMEDLHRHIFWEPDASKLNNENYCRNPDDDAHGPW 420  
QY 421 CYTGNPLIPWDYCPISRCGEGDTPPIV 447  
Db 421 CYTGNPLIPWDYCPISRCGEGDTPPIV 447

RESULT 3  
AAU04275  
ID AAU04275 standard; Peptide; 697 AA.  
XX  
AC AAU04275;  
DT 23-OCT-2001 (first entry)  
XX  
DE Nuclear ligand Pep20 used in nucleic acid transporter system.  
KW Nucleic acid transport; cytolysis; ligand; lysis agent; spacer molecule;  
KW gene therapy; hepatocyte; muscle; bone forming cell.  
XX  
OS Synthetic.  
XX  
FN US6177554-B1.  
XX  
PD 23-JAN-2001.  
XX  
PF 05-JUN-1995; 95US-0462040.  
XX  
PR 14-DEC-1993; 93US-0167641.  
PR 20-MAR-1992; 92US-0855389.  
PR 19-MAR-1993; 93WO-US02725.  
XX  
PA (BAYU ) BAYLOR COLLEGE MEDICINE.  
XX  
PI Woo SLC, Smith LC, Cristiano RJ, Gottchalk S, Sparrow J;  
XX WPI; 2001-365933/38.  
XX  
PT Nucleic acid transport system, useful for creating transgenic animals  
PT for assessing human disease such as cancer in an animal model -  
XX  
PS Disclosure; Fig 23A; 111pp; English.  
XX  
CC The sequence represents the nuclear ligand, Pep20, used in a nucleic acid  
CC transporter system. The nucleic acid transporter system uses nucleic acid  
CC binding complexes containing surface ligands which are capable of binding  
CC to a cell surface receptor and entering the cell through cytolysis. The  
CC compounds of the invention are either ligands, binding molecules (surface  
CC ligands), lysis agents, spacer molecules or their intermediates. The  
CC ligands, binding molecules, lysis agents and spacer molecules are used in  
CC nucleic acid transporter systems to deliver nucleic acid into specific  
CC cells e.g. in gene therapy to deliver nucleic acid into hepatocytes,  
XX muscle cells or bone forming cells.  
SQ Sequence 697 AA;

Query Match 99.9%; Score 2601; DB 22; Length 697;  
Best Local Similarity 99.8%; Pred. No. 1.3e-165;  
Matches 446; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ERKRNTTHEFKSAKTLIKIDPALKIKTKVNTADOCANRCKTRNKGLPFTCKAFVFDK 60  
Db 1 QRKRNTTHEFKSAKTLIKIDPALKIKTKVNTADOCANRCKTRNKGLPFTCKAFVFDK 60  
QY 61 ARKQCLWFPFNSMSGVKKFGEHFDLYENKDYIRNCIIGKRSYKGVTSITKSIKQCP 120  
Db 61 ARKQCLWFPFNSMSGVKKFGEHFDLYENKDYIRNCIIGKRSYKGVTSITKSIKQCP 120  
QY 121 WSSMIPHEHSFLPSSYRGKDLQENYCRNPRGEGPWCFTSNPEVRYEYCDIPQCSEVEC 180

Db 121 WSSMIPHEHSFLPSSYRGKDLQENYCRNPRGEGPWCFTSNPEVRYEYCDIPQCSEVEC 180  
QY 181 MTCNGESYRGLMDHTESGKICQWDHQTPIRHKFLPERYPDKGFDNCRNPDGQPRPW 240  
Db 181 MTCNGESYRGLMDHTESGKICQWDHQTPIRHKFLPERYPDKGFDNCRNPDGQPRPW 240  
QY 241 YTLDPHTRWEYCAIKTCADNTMNDTVPLETTECIQGGEGYRGVTNTIWIINGIPCORWDS 300  
Db 241 YTLDPHTRWEYCAIKTCADNTMNDTVPLETTECIQGGEGYRGVTNTIWIINGIPCORWDS 300  
QY 301 QYPHEHDMTPENFKCKDLRENYCRNPDSGSPWCFTTDPNIRVGYCSQIPNCDSHGGDC 360  
Db 301 QYPHEHDMTPENFKCKDLRENYCRNPDSGSPWCFTTDPNIRVGYCSQIPNCDSHGGDC 360  
QY 361 YRGNGKNTMGNSQTRSGLTCSMDKNMEDLHRHIFWEPDASKLNNENYCRNPDDDAHGPW 420  
Db 361 YRGNGKNTMGNSQTRSGLTCSMDKNMEDLHRHIFWEPDASKLNNENYCRNPDDDAHGPW 420  
QY 421 CYTGNPLIPWDYCPISRCGEGDTPPIV 447  
Db 421 CYTGNPLIPWDYCPISRCGEGDTPPIV 447

RESULT 4  
AAB45838  
ID AAB45838 standard; Protein; 697 AA.  
XX  
AC AAB45838;  
XX  
DT 21-MAR-2001 (first entry)  
XX  
DE Nucleic acid transporter system peptide ligand SEQ ID NO 50.  
DE  
DE  
KW Nucleic acid delivery; nucleic acid transporter system; hormone; enzyme;  
KW growth factor; clotting factor; apolipoprotein; receptor; drug; oncogene;  
KW tumor antigen; tumor suppressor; viral antigen; parasitic antigen;  
KW bacterial antigen.  
XX  
OS Unidentified.  
XX  
PN US6150168-A.  
XX  
PD 21-NOV-2000.  
XX  
PF 05-JUN-1995; 95US-0460971.  
XX  
PR 14-DEC-1993; 93US-0167641.  
PR 20-MAR-1992; 92US-0855389.  
PR 19-MAR-1993; 93WO-US02725.  
XX  
PA (BAYU ) BAYLOR COLLEGE MEDICINE.  
XX  
PI Gottchalk S, Sparrow J, Cristiano RJ, Smith LC, Woo SLC;  
XX WPI; 2001-049093/06.  
XX  
PT Nucleic acid transporter system for delivering nucleic acid into a  
PT cell, useful for delivering proteins and polypeptides to cells,  
PT including growth factors, enzymes, hormones, and tumor suppressors  
XX  
PS Claim 8; Column 115-118; 105pp; English.  
XX  
CC This invention describes a novel system (I) for delivering a nucleic acid  
CC to a cell, comprising a binding complex comprising a ligand binding  
CC molecule noncovalently bound to a nucleic acid and covalently linked to a  
CC surface ligand, and a second binding complex comprising a second binding  
CC molecule noncovalently bound to a nucleic acid and covalently linked to a  
CC nucleic acid ligand. The complexes are simultaneously bound to the nucleic  
CC acid. The nucleic acid transporter system can also be used in a method  
CC for the in vivo targeting of the insertion of DNA into a cell. It can  
CC also be used in processes for producing transformed cell lines. The  
CC system can be used to deliver a variety of proteins and polypeptides,

CC such as hormones, growth factors, clotting factors,  
 CC apolipoproteins, receptors, drugs, oncogenes, tumor antigens, tumor  
 CC suppressors, viral antigens, parasitic antigens, and bacterial antigens.  
 CC The transporter system uses lysis agents to overcome the problems of  
 CC endosomal/lysosomal degradation seen with prior art systems.

XX  
 CC Sequence 697 AA;  
 Query Match 99.9%; Score 2601; DB 22; Length 697;  
 Best Local Similarity 99.8%; Pred. No. 1.3e-165;  
 Matches 446; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ERKRRNTIHEFKSAKTLIKIDPALKIKTKVNTADOCANRCTRNKGLPFTCKAFVDEK 60  
 :|||||  
 Db 1 QKRRNTIHEFKSAKTLIKIDPALKIKTKVNTADOCANRCTRNKGLPFTCKAFVDEK 60  
 QY 61 ARKQCLWFFPNSMSGVKKKEGHEFDLYENKDYIRNCIIIGKRSYKGTVSITKSGIKQCP 120  
 :|||||  
 Db 61 ARKQCLWFFPNSMSGVKKKEGHEFDLYENKDYIRNCIIIGKRSYKGTVSITKSGIKQCP 120  
 QY 121 WSSMIPHEHSFLPSSYRGKDLQENYCRNPRGEGGPGWFTSNPEVRYEVCIDIPOCSEVEC 180  
 :|||||  
 Db 121 WSSMIPHEHSFLPSSYRGKDLQENYCRNPRGEGGPGWFTSNPEVRYEVCIDIPOCSEVEC 180  
 QY 181 MTCNGESYRGLMDHTESGKICORWDHQTTPHKKFLPERYPDKGFDNCRNPDGQPRPWC 240  
 :|||||  
 Db 181 MTCNGESYRGLMDHTESGKICORWDHQTTPHKKFLPERYPDKGFDNCRNPDGQPRPWC 240  
 QY 241 YTLDPHTRWEYCAIKTCADNTMNDTVPLETTETCIGQGEYRGVTNTIWNIGIPQWRWDS 300  
 :|||||  
 Db 241 YTLDPHTRWEYCAIKTCADNTMNDTVPLETTETCIGQGEYRGVTNTIWNIGIPQWRWDS 300  
 QY 301 QYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCTTDPNIRVGYCSQIPNCDMSHGQDC 360  
 :|||||  
 Db 301 QYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCTTDPNIRVGYCSQIPNCDMSHGQDC 360  
 QY 361 YRNGKNYMGNLSTQTRSGLTCSMWKDNMEDLHRHIFWEPDASKLNEYCRNPDHAGHPW 420  
 :|||||  
 Db 361 YRNGKNYMGNLSTQTRSGLTCSMWKDNMEDLHRHIFWEPDASKLNEYCRNPDHAGHPW 420  
 QY 421 CYTGNPLIPWDYCPISRCGEGDTTPTIV 447  
 :|||||  
 Db 421 CYTGNPLIPWDYCPISRCGEGDTTPTIV 447

RESULT 5  
 AAR20005  
 ID AAR20005 standard; Protein; 728 AA.

XX AAR20005;  
 AC  
 DT 24-MAR-1992 (first entry)  
 XX Human hepatocyte growth factor.  
 DE HGF; kidney regeneration; nephritis.

XX Homo sapiens.  
 XX EP462549-A.  
 XX  
 PD 27-DEC-1991.

XX 18-JUN-1991; 91EP-0109923.  
 XX  
 PR 19-JUN-1990; 90JP-0158841.  
 XX (TOYM ) TOYO BOSEKI KK.

XX Nakamuro T;  
 PI  
 DR WPI; 1992-000939/01.  
 DR N-PSDB; AAQ20049.

XX Agent contg. hepatocyte growth factor and carrier - used for  
 PT treating renal diseases and promoting nephrocyte growth and as  
 PT diagnostic for renal diseases  
 XX  
 PS Disclosure; Fig 1; 15pp; English.  
 XX  
 CC Human HGF comprises an alpha-chain of 440 amino acids and a  
 CC beta-chain of 234 amino acids. There are 4 kringle domains in the  
 CC alpha-chain, similar to that of plasmin; the beta-chain has about  
 CC 37 per cent homology with the beta-chain of plasmin having serine  
 CC protease activity. Homology of the amino acid sequence of rat HGF  
 CC and human HGF is 91.6 per cent in the alpha-chain and 88.9 per cent  
 CC in the beta-chain. HGF has been found to be an agent for nephrocyte  
 CC growth and is useful as a treatment for renal failure.

XX Sequence 728 AA;  
 Query Match 99.9%; Score 2601; DB 13; Length 728;  
 Best Local Similarity 99.8%; Pred. No. 1.4e-165;  
 Matches 446; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ERKRRNTIHEFKSAKTLIKIDPALKIKTKVNTADOCANRCTRNKGLPFTCKAFVDEK 60  
 :|||||  
 Db 32 QKRRNTIHEFKSAKTLIKIDPALKIKTKVNTADOCANRCTRNKGLPFTCKAFVDEK 91  
 QY 61 ARKQCLWFFPNSMSGVKKKEGHEFDLYENKDYIRNCIIIGKRSYKGTVSITKSGIKQCP 120  
 :|||||  
 Db 92 ARKQCLWFFPNSMSGVKKKEGHEFDLYENKDYIRNCIIIGKRSYKGTVSITKSGIKQCP 151  
 QY 121 WSSMIPHEHSFLPSSYRGKDLQENYCRNPRGEGGPGWFTSNPEVRYEVCIDIPOCSEVEC 180  
 :|||||  
 Db 152 WSSMIPHEHSFLPSSYRGKDLQENYCRNPRGEGGPGWFTSNPEVRYEVCIDIPOCSEVEC 211  
 QY 181 MTCNGESYRGLMDHTESGKICORWDHQTTPHKKFLPERYPDKGFDNCRNPDGQPRPWC 240  
 :|||||  
 Db 212 MTCNGESYRGLMDHTESGKICORWDHQTTPHKKFLPERYPDKGFDNCRNPDGQPRPWC 271  
 QY 241 YTLDPHTRWEYCAIKTCADNTMNDTVPLETTETCIGQGEYRGVTNTIWNIGIPQWRWDS 300  
 :|||||  
 Db 272 YTLDPHTRWEYCAIKTCADNTMNDTVPLETTETCIGQGEYRGVTNTIWNIGIPQWRWDS 331  
 QY 301 QYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCTTDPNIRVGYCSQIPNCDMSHGQDC 360  
 :|||||  
 Db 332 QYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCTTDPNIRVGYCSQIPNCDMSHGQDC 391  
 QY 361 YRNGKNYMGNLSTQTRSGLTCSMWKDNMEDLHRHIFWEPDASKLNEYCRNPDHAGHPW 420  
 :|||||  
 Db 392 YRNGKNYMGNLSTQTRSGLTCSMWKDNMEDLHRHIFWEPDASKLNEYCRNPDHAGHPW 451  
 QY 421 CYTGNPLIPWDYCPISRCGEGDTTPTIV 447  
 :|||||  
 Db 452 CYTGNPLIPWDYCPISRCGEGDTTPTIV 478

RESULT 6  
 AAR25676  
 ID AAR25676 standard; protein; 728 AA.

XX AAR25676;  
 AC  
 DT 20-JAN-1993 (first entry)  
 XX Recombinant human hepatocyte growth factor.

XX HGF; enhance growth; preparing transgenic animals; hepatic disease;  
 KW clinical diagnostic reagent; drug.

XX Homo sapiens.  
 XX JP04183394-A.  
 XX  
 PD 30-JUN-1992.

```

XX PF 19-NOV-1990; 90JP-0314548.
XX PR 19-NOV-1990; 90JP-0314548.
XX PA (NAKA/) NAKAMURA T.
XX PA (TOYM ) TOYOKO KK.
XX DR WPI; 1992-265591/32.
XX Recombinant human hepatocyte growth factor and DNA encoding it -
XX useful for diagnosis and treatment of hepatic disease and
XX PT transgenic animal prepn.
XX PS Disclosure; Page 11; 28pp; Japanese.
XX CC This sequence represents a recombinant human hepatocyte growth
XX factor. It has physiological activity, and using it enhanced growth
XX of hepatocytes is possible. It is useful as a clinical diagnostic
XX reagent, or a drug for treating hepatic disease.
XX CC See also AAR25676-92, AAQ26713-27.
XX SQ Sequence 728 AA;

Query Match 99.9%; Score 2601; DB 13; Length 728;
Best Local Similarity 99.8%; Pred. No. 1.4e-165;
Matches 446; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ERKRRNTIHEFKSAKTLTIKIDPALKIKTKKVNADQCANRCKTRNKGLPFTCKAFVFDK 60
Db :|||||
Qy 32 QKRRNTIHEFKSAKTLTIKIDPALKIKTKKVNADQCANRCKTRNKGLPFTCKAFVFDK 91
Db :|||||
Qy 61 ARKQCLWFFPNSMSGVKKFEGHEFDLYENKDYIRNCIIIGKRSYKGTVSIITKSGIKCOP 120
Db :|||||
Qy 92 ARKQCLWFFPNSMSGVKKFEGHEFDLYENKDYIRNCIIIGKRSYKGTVSIITKSGIKCOP 151
Db :|||||
Qy 121 WSSMIPHEHSFLPSSYRGKDLQENYCRNPRGEGGPGWCFSTSNPEVRYEVCIDIPOQSEVEC 180
Db :|||||
Qy 152 WSSMIPHEHSFLPSSYRGKDLQENYCRNPRGEGGPGWCFSTSNPEVRYEVCIDIPOQSEVEC 211
Db :|||||
Qy 181 MTCNGESYRGLMDHTESGKICORWDHQTPIRHKFLPERYPDKGFDNNYCRNPDGQPRPWC 240
Db :|||||
Qy 212 MTCNGESYRGLMDHTESGKICORWDHQTPIRHKFLPERYPDKGFDNNYCRNPDGQPRPWC 271
Db :|||||
Qy 241 YTLDPHTRWEYCAIKTCADNTMNDTVPLETTECIQGOGEGYGTVTNTIWNIGIPQQRWDS 300
Db :|||||
Qy 272 YTLDPHTRWEYCAIKTCADNTMNDTVPLETTECIQGOGEGYGTVTNTIWNIGIPQQRWDS 331
Db :|||||
Qy 301 QYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDMSHGQDC 360
Db :|||||
Qy 332 QYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDMSHGQDC 391
Db :|||||
Qy 361 YRGNKNTMGNLSQTRSGLTCSMDKNMEDLHRHIFWEPDASKLNYCRNPDDDAHGPW 420
Db :|||||
Qy 392 YRGNKNTMGNLSQTRSGLTCSMDKNMEDLHRHIFWEPDASKLNYCRNPDDDAHGPW 451
Db :|||||
Qy 421 CYTGNPLIPWDYCPISRCGDDTPTIV 447
Db :|||||
Qy 452 CYTGNPLIPWDYCPISRCGDDTPTIV 478
Db :|||||

RESULT 7
AAR40862
ID AAR40862 standard; Protein; 728 AA.
XX AC AAR40862;
XX DF 14-MAR-1994 (first entry)
XX DE Competative inhibitor of HGF.
XX KW HGF; hepatocyte growth factor; transformation; antagonist;
XX liver disease.

```

```

XX OS Homo sapiens.
XX FH Key
XX FT Peptide
XX FT 1..31
XX FT /label= sig_peptide
XX FT 32..728
XX FT /label= mat_protein
XX FT Misc-difference 494
XX FT /note= "site of possible mutation"
XX XX
XX PN JP05208998-A.
XX PD 20-AUG-1993.
XX PF 25-DEC-1991; 91JP-0357040.
XX PR 25-DEC-1991; 91JP-0357040.
XX PA (MITU ) MITSUBISHI KASEI CORP.
XX DR WPI; 1993-297806/38.
XX DR N-PSDB; AAQ47832.
XX PT New protein is competitive inhibitor of hepatocyte growth factor
XX PT - used in animal model of liver disease
XX PS Claim 7; Page 8-11; 20pp; Japanese.
XX CC The gene encodes an inhibitor of HGF, which has a molecular weight
XX of ca. 76000 to 92000 as determined by SDS-PAGE. Human modified
XX (NP) HGF-A and -B fragments were prepared from human HGF cDNA.
XX CC This was amplified by PCR to give a human NP-HGF DNA fragment that
XX CC was transformed into E.coli. (see AAQ47833) for a related sequence.
XX SQ Sequence 728 AA;

Query Match 99.9%; Score 2601; DB 14; Length 728;
Best Local Similarity 99.8%; Pred. No. 1.4e-165;
Matches 446; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ERKRRNTIHEFKSAKTLTIKIDPALKIKTKKVNADQCANRCKTRNKGLPFTCKAFVFDK 60
Db :|||||
Qy 32 QKRRNTIHEFKSAKTLTIKIDPALKIKTKKVNADQCANRCKTRNKGLPFTCKAFVFDK 91
Db :|||||
Qy 61 ARKQCLWFFPNSMSGVKKFEGHEFDLYENKDYIRNCIIIGKRSYKGTVSIITKSGIKCOP 120
Db :|||||
Qy 92 ARKQCLWFFPNSMSGVKKFEGHEFDLYENKDYIRNCIIIGKRSYKGTVSIITKSGIKCOP 151
Db :|||||
Qy 121 WSSMIPHEHSFLPSSYRGKDLQENYCRNPRGEGGPGWCFSTSNPEVRYEVCIDIPOQSEVEC 180
Db :|||||
Qy 152 WSSMIPHEHSFLPSSYRGKDLQENYCRNPRGEGGPGWCFSTSNPEVRYEVCIDIPOQSEVEC 211
Db :|||||
Qy 181 MTCNGESYRGLMDHTESGKICORWDHQTPIRHKFLPERYPDKGFDNNYCRNPDGQPRPWC 240
Db :|||||
Qy 212 MTCNGESYRGLMDHTESGKICORWDHQTPIRHKFLPERYPDKGFDNNYCRNPDGQPRPWC 271
Db :|||||
Qy 241 YTLDPHTRWEYCAIKTCADNTMNDTVPLETTECIQGOGEGYGTVTNTIWNIGIPQQRWDS 300
Db :|||||
Qy 272 YTLDPHTRWEYCAIKTCADNTMNDTVPLETTECIQGOGEGYGTVTNTIWNIGIPQQRWDS 331
Db :|||||
Qy 301 QYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDMSHGQDC 360
Db :|||||
Qy 332 QYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDMSHGQDC 391
Db :|||||
Qy 361 YRGNKNTMGNLSQTRSGLTCSMDKNMEDLHRHIFWEPDASKLNYCRNPDDDAHGPW 420
Db :|||||
Qy 392 YRGNKNTMGNLSQTRSGLTCSMDKNMEDLHRHIFWEPDASKLNYCRNPDDDAHGPW 451
Db :|||||
Qy 421 CYTGNPLIPWDYCPISRCGDDTPTIV 447
Db :|||||
Qy 452 CYTGNPLIPWDYCPISRCGDDTPTIV 478
Db :|||||

```

```

RESULT 8
AAR40863
ID AAR40863 standard; Protein; 728 AA.
XX
AC AAR40863;
XX
DT 14-MAR-1994 (first entry)
XX
DE Competitive inhibitor of HGF.
XX
KW HGF; hepatocyte growth factor; transformation; antagonist;
liver disease.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Peptide 1..31
FT Protein 32..728
FT /label= sig_peptide
FT /label= mat_protein
FT Misc-difference 489
FT /note= "site of possible mutation"
FT Misc-difference 491
FT /note= "site of possible mutation"
FT Misc-difference 494
FT /note= "site of possible mutation"
XX
XX JP05208998-A.
XX
XX 20-AUG-1993.
XX
XX 25-DEC-1991; 91JP-0357040.
XX
XX 25-DEC-1991; 91JP-0357040.
XX
XX (MITU ) MITSUBISHI KASEI CORP.
XX
XX WPI; 1993-297806/38.
XX
XX N-PSDB; AAQ47833.
XX
XX New protein is competitive inhibitor of hepatocyte growth factor
XX - used in animal model of liver disease
XX
XX Claim 6; Page 11-14; 20pp; Japanese.
XX
XX The gene encodes an inhibitor of HGF, which has a molecular weight
XX of ca. 76000 to 92000 as determined by SDS-PAGE. Human modified
XX (NP) HGF-A and -B fragments were prepared from human HGF cDNA.
XX This was amplified by PCR to give a human NP-HGF DNA fragment that
XX was transformed into E.coli. (see AAQ47833) for a related sequence.
XX
XX Sequence 728 AA;
XX
Query Match 99.9%; Score 2601; DB 14; Length 728;
Best Local Similarity 99.8%; Pred. No. 1.4e-165;
Matches 446; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 ERKRRNTTHERFKSAKTLIKIDPALKTKTKVNTADQCANRCTRNKGLPFTCKAFVFDK 60
DB 32 ORKRRNTTHERFKSAKTLIKIDPALKTKTKVNTADQCANRCTRNKGLPFTCKAFVFDK 91
QY 61 ARKQCLWFPNFMSSGVKKEGHEFDLYENKDYIRNCIIIGKRSYKGTSTVITKSGIKQCP 120
DB 92 ARKQCLWFPNFMSSGVKKEGHEFDLYENKDYIRNCIIIGKRSYKGTSTVITKSGIKQCP 151
QY 121 WSSMTPHEHSFLPSSYRGKDLOENYCRNPRGEGGPGWFTSNPEVRYEVCDIPQCSVEEC 180
DB 152 WSSMTPHEHSFLPSSYRGKDLOENYCRNPRGEGGPGWFTSNPEVRYEVCDIPQCSVEEC 211
QY 181 MTCNGESYRGLMDHTESGKICQWRDHQTPHRRHKFLPERYPDKGFDNDYCRNPDQGPRAW 240
DB 212 MTCNGESYRGLMDHTESGKICQWRDHQTPHRRHKFLPERYPDKGFDNDYCRNPDQGPRAW 271

```

Qy 61 ARKQCLWFPNFMSSGVKKERGFHEFDLYENKDYIRNCIIIGKGRSYKGVTSITKSGIKQCP 120  
Db 92 ARKQCLWFPNFMSSGVKKERGFHEFDLYENKDYIRNCIIIGKGRSYKGVTSITKSGIKQCP 151  
Qy 121 WSSMIPHEHSFLPSSYRGKDLQENYCRNPRGEGGPGWFTSNPEVRYEVCIDIPOCSEVEC 180  
Db 152 WSSMIPHEHSFLPSSYRGKDLQENYCRNPRGEGGPGWFTSNPEVRYEVCIDIPOCSEVEC 211  
Qy 181 MTCNGESYRGLMDHTEGSKICQWRDHTQPHRHKFLPERYPDKGFDNNYCRNPDGQPRWC 240  
Db 212 MTCNGESYRGLMDHTEGSKICQWRDHTQPHRHKFLPERYPDKGFDNNYCRNPDGQPRWC 271  
Qy 241 YLDPHTRWEYCAIKTCADNTMNDTVPLETTECIQOGEGYRGTVNTIWNIGPCQRWDS 300  
Db 272 YLDPHTRWEYCAIKTCADNTMNDTVPLETTECIQOGEGYRGTVNTIWNIGPCQRWDS 331  
Qy 301 QYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDMSHGQDC 360  
Db 332 QYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDMSHGQDC 391  
Qy 361 YRGNGKNTYMGNLSQTRSGLTCSMDKNMEDLHRHIFWEPDASKNLNENYCRNPDGDAHGWP 420  
Db 392 YRGNGKNTYMGNLSQTRSGLTCSMDKNMEDLHRHIFWEPDASKNLNENYCRNPDGDAHGWP 451  
Qy 421 CYTGNPLIPWDYCPISRCGEGDTTPTIV 447  
Db 452 CYTGNPLIPWDYCPISRCGEGDTTPTIV 478

RESULT 10  
AAW00338

ID AAW00338 standard; protein: 728 AA.  
AC AAW00338;  
XX 27-JUN-1997 (first entry)  
XX Human hepatic parenchymal cell growth factor.  
XX Human; hepatic; parenchymal cell; growth factor; HGF; treatment;  
KW prevention; ischaemia; ischaemic; disease; reperfusion; disorder;  
KW blood; liver; transplant; acute; failure; ischemia; ischemic;  
KW organ.  
XX Homo sapiens.  
OS WO9632960-A1.  
PN 24-OCT-1996.  
XX 19-APR-1996; 96WO-JP01065.  
XX 21-APR-1995; 95JP-0096994.  
XX (MITU ) MITSUBISHI CHEM CORP.  
XX Gamba M, Yonehana T;  
XX WPI; 1996-485555/48.  
XX Hepatic parenchymal cell growth factors - used as remedy and/or  
PT preventive medicine for ischaemic diseases  
XX Claim 9; Pages 7-10; 16pp; Japanese.  
XX The present sequence is the human hepatic parenchymal cell growth  
CC factor (HGF), which has a molecular weight of 76-92 kD estimated by  
CC SDS-PAGE, parenchymal cell growth activity, which is lost following  
CC heat treatment at 80 degrees C for 10 minutes and strong affinity  
CC towards heparin. It may be used to treat and prevent ischaemic  
CC diseases, e.g. blood reperfusion disorder, ischaemic reperfusion  
CC disorders in liver transplant, acute liver failure and organ  
CC disorder during organ transplant. For adults, it is administered at

CC a level of 1 microg to 10 mg/kg/day, preferably 10-1000  
CC microg/kg/day.  
CC A pig liver derived cultured epithelial cell line was cultured in  
CC DMEM, cow embryo serum and N-2-hydroxyethylpiperazine ethane  
CC sulphonic acid to confluence. The medium was then changed to serum  
CC and glucose free DMEM, and the oxygen concentration reduced to less  
CC than 2%. The cells were cultured for 6 hours, and then for another  
CC hour at 95% air/5% carbon dioxide (reoxxygenation). HGF was added  
CC just before hypoxiation or reoxxygenation. HGF suppressed free  
CC lactic acid dehydrogenase (LDH) by 49.6% when added before  
CC hypoxiation, and by 60.9% when added before reoxxygenation, at 50  
CC ng/ml (free LDL was 100% in the absence of HGF).  
XX  
SQ Sequence 728 AA;  
Query Match 99.9%; Score 2601; DB 17; Length 728;  
Best Local Similarity 99.8%; Pred. No. 1.4e-165;  
Matches 446; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 ERKRNTIHEFKSAKTTLLIKIDPALKIKTKKVNADOCANRCTRNGKLPFTCKAFVFDK 60  
Db 32 QKRKRNTIHEFKSAKTTLLIKIDPALKIKTKKVNADOCANRCTRNGKLPFTCKAFVFDK 91  
Qy 61 ARKQCLWFPNFMSSGVKKERGFHEFDLYENKDYIRNCIIIGKGRSYKGVTSITKSGIKQCP 120  
Db 92 ARKQCLWFPNFMSSGVKKERGFHEFDLYENKDYIRNCIIIGKGRSYKGVTSITKSGIKQCP 151  
Qy 121 WSSMIPHEHSFLPSSYRGKDLQENYCRNPRGEGGPGWFTSNPEVRYEVCIDIPOCSEVEC 180  
Db 152 WSSMIPHEHSFLPSSYRGKDLQENYCRNPRGEGGPGWFTSNPEVRYEVCIDIPOCSEVEC 211  
Qy 181 MTCNGESYRGLMDHTEGSKICQWRDHTQPHRHKFLPERYPDKGFDNNYCRNPDGQPRWC 240  
Db 212 MTCNGESYRGLMDHTEGSKICQWRDHTQPHRHKFLPERYPDKGFDNNYCRNPDGQPRWC 271  
Qy 241 YLDPHTRWEYCAIKTCADNTMNDTVPLETTECIQOGEGYRGTVNTIWNIGPCQRWDS 300  
Db 272 YLDPHTRWEYCAIKTCADNTMNDTVPLETTECIQOGEGYRGTVNTIWNIGPCQRWDS 331  
Qy 301 QYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDMSHGQDC 360  
Db 332 QYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDMSHGQDC 391  
Qy 361 YRGNGKNTYMGNLSQTRSGLTCSMDKNMEDLHRHIFWEPDASKNLNENYCRNPDGDAHGWP 420  
Db 392 YRGNGKNTYMGNLSQTRSGLTCSMDKNMEDLHRHIFWEPDASKNLNENYCRNPDGDAHGWP 451  
Qy 421 CYTGNPLIPWDYCPISRCGEGDTTPTIV 447  
Db 452 CYTGNPLIPWDYCPISRCGEGDTTPTIV 478  
RESULT 11  
AAW00340  
ID AAW00340 standard; protein: 728 AA.  
XX  
AC AAW00340;  
XX 09-DEC-1996 (first entry)  
XX Wild type hepatocyte growth factor.  
DE Human; hepatocyte growth factor; HGF; huHGF; serum; proteolytic cleavage;  
XX pro-hormone; beta subunit; alpha subunit; kringle domain; prothrombin;  
KW plasminogen; catalytic domain; serine protease; HGF variant;  
KW HGF receptor; malignancy; chronic HGF receptor activation.  
XX Homo sapiens.  
OS  
XX Key Location/Qualifiers  
FH Cleavage-site 494..495  
FT /label= Proteolytic\_cleavage\_site  
FT /note= "generates alpha and beta subunits"

FT Disulfide-bond 487..604

FT Peptide 1..54

FT Peptide /note= "prepro-sequence"

FT Peptide 1..31

FT Domain /note= "Hydrophobic signal peptide"

FT Domain 128..206

FT Domain /label= Kringle\_1\_domain

FT Domain 211..288

FT Domain /label= Kringle\_2\_domain

FT Domain 303..383

FT Domain /label= Kringle\_3\_domain

FT Domain 391..464

FT Domain /label= Kringle\_4\_domain

FT Modified-site 294

FT Modified-site /label= N-linked-glycosylation\_site

FT Modified-site 402

FT Modified-site /label= N-linked-glycosylation\_site

FT Modified-site 566

FT Modified-site /label= N-linked-glycosylation\_site

FT Modified-site 653

FT Modified-site /label= N-linked-glycosylation\_site

FT Misc-difference 494

FT /note= "Position of opt. substitution, esp. Glu, Asp or Asn"

FT Misc-difference 495

FT /note= "Position of opt. substitution, esp. Tyr or Phe"

FT Misc-difference 534

FT /note= "Position of opt. substitution"

FT Misc-difference 673

FT /note= "Position of opt. substitution"

FT Misc-difference 692

FT /note= "Position of opt. substitution"

FT FT

XX US5547856-A.

XX 20-AUG-1996.

XX 18-MAY-1992; 92US-0884811.

XX 13-JUL-1993; 93US-0087783.

XX 18-MAY-1992; 92US-0884811.

XX 18-MAY-1992; 92US-0885971.

XX (GETH ) GENENTECH INC.

XX Godowski PJ, Lokker NA, Mark MR;

XX WPI; 1996-392634/39.

XX New hepatocyte growth factor variants - are resistant to in vivo

XX proteolytic cleavage into a 2-chain form, useful as HGF antagonists

XX Disclosure; Column 51-56; 39pp; English.

XX This sequence represents the wild type full length sequence of human

XX hepatocyte growth factor (HGF). HGF is isolated from human serum

XX and is a disulphide linked heterodimer derived by proteolytic cleavage

XX of this pro-hormone form between residues 494 and 495. This generates a

XX molecule composed of an alpha-subunit of 440 amino acids (mol. wt. 69

XX kD) and a beta-subunit of 234 amino acids (mol. wt. 34 kD). The alpha

XX and beta subunits are encoded by a single open reading frame. The alpha

XX subunit contains four kringle domains based on their homology to

XX kringle-like domains in other proteins, e.g. prothrombin, plasminogen.

XX The beta subunit shows high homology to the catalytic domain of serine

XX proteases. However two of the three residues which form the catalytic

XX triad of serine proteases are not conserved in HGF. Therefore, the

XX precise function of the beta chain remains unknown. The invention

XX includes HGF variants which retain HGF receptor binding activity without

XX having the biological activity of wild-type HGF. They can be used for

XX the treatment of pathological conditions associated with the activation

XX of a HGF receptor such as malignancies associated with chronic HGF

XX receptor activation.

SQ	Sequence	728 AA;
Query Match	99.9%; Score 2601; DB 17; Length 728;	
Best Local Similarity	99.8%; Pred. No. 1.4e-165;	
Matches 446; Conservative	1; Mismatches 0; Indels 0; Gaps 0;	
QY	1 ERKRRNTIHEFKSAKTTLLIKIDPALKIKTKVNTADQCANRCKTRNKGLPFTCKAEVFK 60	
DB	32 QRRKRRNTIHEFKSAKTTLLIKIDPALKIKTKVNTADQCANRCKTRNKGLPFTCKAEVFK 91	
QY	61 ARKOCLEWFFPNSMSSGVKKKEFGHEFDLYENKDYIRNCIIIGKGRSYKGTYSITKSGIKCOP 120	
DB	92 ARKOCLEWFFPNSMSSGVKKKEFGHEFDLYENKDYIRNCIIIGKGRSYKGTYSITKSGIKCOP 151	
QY	121 WSSMTPHEHSLFSPSSYRGKDLQENYCRNPRGEGGPMCFSTSNPEVRYEVCIDIQCSEVEC 180	
DB	152 WSSMTPHEHSLFSPSSYRGKDLQENYCRNPRGEGGPMCFSTSNPEVRYEVCIDIQCSEVEC 211	
QY	181 MTCNGESYRGLMDHTESGKTCQRDWHOTPHRHKFLPERYPDKGFDNNYCRNPDGQPRPW 240	
DB	212 MTCNGESYRGLMDHTESGKTCQRDWHOTPHRHKFLPERYPDKGFDNNYCRNPDGQPRPW 271	
QY	241 YLDPHTRWEYCAIKTCADNTMNDTDVPLETTECICQGGEGYRGTVNTIWNIGPCQRWDS 300	
DB	272 YLDPHTRWEYCAIKTCADNTMNDTDVPLETTECICQGGEGYRGTVNTIWNIGPCQRWDS 331	
QY	301 QYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDMSHGQDC 360	
DB	332 QYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDMSHGQDC 391	
QY	361 YRGNGKNTYMGNLSQTRSGLTCSMDKKNMEDLHRHIFWEPDASKLNENYCRNPDGDAHGPW 420	
DB	392 YRGNGKNTYMGNLSQTRSGLTCSMDKKNMEDLHRHIFWEPDASKLNENYCRNPDGDAHGPW 451	
QY	421 CYTGNPLIPWDYCPISRCGEGDTTPTIV 447	
DB	452 CYTGNPLIPWDYCPISRCGEGDTTPTIV 478	
RESULT 12		
AAR87522		
ID	AAR87522 standard; protein; 728 AA.	
XX		
AC	AAR87522;	
XX		
DT	21-JUN-1996 (first entry)	
XX		
DE	Mutant hepatocyte growth factor #1 with changes at residues 491-495.	
XX		
KW	Hepatocyte growth factor; protease; cleavage recognition site; Factor Xa;	
KW	injury; blood vessel; point mutation.	
OS	Synthetic.	
XX		
XX	Key Location/Qualifiers	
FT	Misc-difference 491..495	
FT	/note= "change from wild type sequence:	
FT	Lys-Glu-Leu-Arg-Val to Ile-Glu-Gly-Arg-Thr"	
XX		
PN	JP07304796-A.	
XX		
PD	21-NOV-1995.	
XX		
PF	07-MAY-1994; 94JP-0117506.	
XX		
PR	07-MAY-1994; 94JP-0117506.	
XX		
PA	(TERU ) TERUMO CORP.	
XX		
XX	WPI; 1996-035890/04.	
XX		
PT	Novel derivative of hepatocyte growth factor - comprises a mutation	
PT	at positions 492-494, for use in treating injured blood vessels	

```
XX PS Disclosure: Page -: 8pp; Japanese.
XX CC The amino acid of the hepatocyte growth factor (HGF) generated to
CC contain amino acid changes at residues 492-494. (Note - this sequence is
CC not given in the specification but is based on the HGF sequence given in
CC AAR20005). The generated mutant contains the sequence
CC Ile-Glu-Gly-Thr, which replaces the wild type sequence
CC Lys-Glu-Leu-Arg-Val. The HGF is translated as single chain protein which
CC is activated to a two chain protein by protease cleavage between residues
CC 494-5, producing the alpha and beta chains. The novel sequence alters the
CC protease recognition site, putatively introducing a Factor Xa recognition
CC site. The novel HGF deriv. can be activated specifically at the site of
CC injury in a blood vessel. (See AAR87522-5 for different HGF derivs.). The
CC novel HGF derivs. were generated by point mutations using the
CC oligonucleotides AAT06762-5.
XX SQ Sequence 728 AA;

Query Match 99.9%; Score 2601; DB 17; Length 728;
Best Local Similarity 99.8%; Pred. No. 1.4e-165;
Matches 446; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ERKRNTTHERFKSAKTLIKIDPALKIKTKKVNADQCANRCKTRNKGLPFTCKAFVFDK 60
Db :|||||
Qy 32 QKRKNTTHERFKSAKTLIKIDPALKIKTKKVNADQCANRCKTRNKGLPFTCKAFVFDK 91
Db :|||||
Qy 61 ARKQCLWFPFNSMSGVKKFEGHEFDLYENKDYIRNCIIIGKGRSYKGTVSIITKSGIKCQP 120
Db :|||||
Qy 92 ARKQCLWFPFNSMSGVKKFEGHEFDLYENKDYIRNCIIIGKGRSYKGTVSIITKSGIKCQP 151
Db :|||||
Qy 121 WSSMTPHEHSFLPSSYRGKDLQENYCRNPRGEGGPGWCFTSNPEVRYEVCIDIPOCSEVEC 180
Db :|||||
Qy 152 WSSMTPHEHSFLPSSYRGKDLQENYCRNPRGEGGPGWCFTSNPEVRYEVCIDIPOCSEVEC 211
Db :|||||
Qy 181 MTCNGESYRGLMDHTESGKICORWDHQPTRHKFLPERYPDKGFDNYCRNPDGQPRPWC 240
Db :|||||
Qy 212 MTCNGESYRGLMDHTESGKICORWDHQPTRHKFLPERYPDKGFDNYCRNPDGQPRPWC 271
Db :|||||
Qy 241 YTLDPHTREWECAIKTCADNTMNDTDVPLETTECIGOGEGYRGVTNTWNGIPCORWDS 300
Db :|||||
Qy 272 YTLDPHTREWECAIKTCADNTMNDTDVPLETTECIGOGEGYRGVTNTWNGIPCORWDS 331
Db :|||||
Qy 301 QYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDMSHGQDC 360
Db :|||||
Qy 332 QYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDMSHGQDC 391
Db :|||||
Qy 361 YRGNGKNTMGNLSQTRSGLTCSMWDKNMEDLHRHFWEPDASKLNEYCRNPDGDDAHGPW 420
Db :|||||
Qy 392 YRGNGKNTMGNLSQTRSGLTCSMWDKNMEDLHRHFWEPDASKLNEYCRNPDGDDAHGPW 451
Db :|||||
Qy 421 CVTGNPLIPWDYCPISRCGEGDTTPTIV 447
Db :|||||
Qy 452 CVTGNPLIPWDYCPISRCGEGDTTPTIV 478
Db :|||||

RESULT 13
AAR87523
ID AAR87523 standard; protein; 728 AA.
XX AC AAR87523;
XX DT 21-JUN-1996 (first entry)
XX DE Mutant hepatocyte growth factor #2 with changes at residues 491-495.
XX KW Hepatocyte growth factor; protease; cleavage recognition site; Factor Xa;
XX OS injury; blood vessel; point mutation.
XX OS Synthetic.
XX FH key Location/Qualifiers
XX FT Misc-difference 491..495
```

```
FT FT /note= "change from wild type sequence:
XX XX Lys-Glu-Leu-Arg-Val to Lys-Gln-Gly-Arg-Ile"
PN JP07304796-A.
XX 21-NOV-1995.
XX 07-MAY-1994; 94JP-0117506.
XX 07-MAY-1994; 94JP-0117506.
XX (TERU ) TERUMO CORP.
XX WPI; 1996-035890/04.
XX PT Novel derivative of hepatocyte growth factor - comprises a mutation
XX at positions 492-494, for use in treating injured blood vessels
XX PS Disclosure: Page -: 8pp; Japanese.
XX CC The amino acid of the hepatocyte growth factor (HGF) generated to
XX contain amino acid changes at residues 492-494. (Note - this sequence is
XX not given in the specification but is based on the HGF sequence given in
XX AAR20005). The generated mutant contains the sequence
XX Lys-Glu-Gly-Arg-Val. The HGF is translated as single chain protein which
XX is activated to a two chain protein by protease cleavage between residues
XX 494-5, producing the alpha and beta chains. The novel sequence alters the
XX protease recognition site, putatively introducing a Factor Xa recognition
XX site. The novel HGF deriv. can be activated specifically at the site of
XX injury in a blood vessel. (See AAR87522-5 for different HGF derivs.). The
XX novel HGF derivs. were generated by point mutations using the
XX oligonucleotides AAT06762-5.
XX SQ Sequence 728 AA;

Query Match 99.9%; Score 2601; DB 17; Length 728;
Best Local Similarity 99.8%; Pred. No. 1.4e-165;
Matches 446; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ERKRNTTHERFKSAKTLIKIDPALKIKTKKVNADQCANRCKTRNKGLPFTCKAFVFDK 60
Db :|||||
Qy 32 QKRKNTTHERFKSAKTLIKIDPALKIKTKKVNADQCANRCKTRNKGLPFTCKAFVFDK 91
Db :|||||
Qy 61 ARKQCLWFPFNSMSGVKKFEGHEFDLYENKDYIRNCIIIGKGRSYKGTVSIITKSGIKCQP 120
Db :|||||
Qy 92 ARKQCLWFPFNSMSGVKKFEGHEFDLYENKDYIRNCIIIGKGRSYKGTVSIITKSGIKCQP 151
Db :|||||
Qy 121 WSSMTPHEHSFLPSSYRGKDLQENYCRNPRGEGGPGWCFTSNPEVRYEVCIDIPOCSEVEC 180
Db :|||||
Qy 152 WSSMTPHEHSFLPSSYRGKDLQENYCRNPRGEGGPGWCFTSNPEVRYEVCIDIPOCSEVEC 211
Db :|||||
Qy 181 MTCNGESYRGLMDHTESGKICORWDHQPTRHKFLPERYPDKGFDNYCRNPDGQPRPWC 240
Db :|||||
Qy 212 MTCNGESYRGLMDHTESGKICORWDHQPTRHKFLPERYPDKGFDNYCRNPDGQPRPWC 271
Db :|||||
Qy 241 YTLDPHTREWECAIKTCADNTMNDTDVPLETTECIGOGEGYRGVTNTWNGIPCORWDS 300
Db :|||||
Qy 272 YTLDPHTREWECAIKTCADNTMNDTDVPLETTECIGOGEGYRGVTNTWNGIPCORWDS 331
Db :|||||
Qy 301 QYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDMSHGQDC 360
Db :|||||
Qy 332 QYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDMSHGQDC 391
Db :|||||
Qy 361 YRGNGKNTMGNLSQTRSGLTCSMWDKNMEDLHRHFWEPDASKLNEYCRNPDGDDAHGPW 420
Db :|||||
Qy 392 YRGNGKNTMGNLSQTRSGLTCSMWDKNMEDLHRHFWEPDASKLNEYCRNPDGDDAHGPW 451
Db :|||||
Qy 421 CVTGNPLIPWDYCPISRCGEGDTTPTIV 447
Db :|||||
Qy 452 CVTGNPLIPWDYCPISRCGEGDTTPTIV 478
Db :|||||
```

[illegible]

FT Misc-difference 491..495  
FT /note= "change from wild type sequence:  
FT Iys-Glu-Ieu-Arg-val to Iys-Gln-Glv-Arg-Thr"  
FT

XX	JP07304796-A.	
PN		
XX		
XX	21-NOV-1995.	
XX		
XX	07-MAY-1994; 94JP-0117506.	
XX		
XX	07-MAY-1994; 94JP-0117506.	
XX		
XX	(TERU ) TERUMO CORP.	
PA		
XX	WPI: 1996-035890/04.	
DR		
XX		
XX	Novel derivative of hepatocyte growth factor - comprises a mutation	
PT	at positions 492-494, for use in treating injured blood vessels	
PT		
XX		
XX	disclosure: Page -: 8pp: Japanese.	
ps		

xx The amino acid of the hepatocyte growth factor (HGF) generated to  
CC contain amino acid changes at residues 492-494. (Note - this sequence is  
CC not given in the specification but is based on the HGF sequence given in  
CC AAR20005). The generated mutant contains the sequence  
CC Lys-Gln-Glu-Arg-Thr, which replaces the wild type sequence  
CC Lys-Glu-Leu-Arg-Val. The HGF is translated as single chain protein which  
CC is activated to a two chain protein by protease cleavage between residues  
CC 494-5, producing the alpha and beta chains. The novel sequence alters the  
CC protease recognition site, putatively introducing a Factor Xa recognition  
CC site. The novel HGF deriv. can be activated specifically at the site of  
CC injury in a blood vessel. (See AAR87522-5 for different HGF derivs.). The  
CC novel HGF derivs. were generated by point mutations using the  
CC oligonucleotides AAT06762-5.

[illegible]

Query Match 99.9%; Score 2601; DB 17; Length 728;  
Best Local Similarity 99.8%; Pred. No. 1.4e-165;  
Matches 446; Conservative 1; Mismatches 0; Indels 0; Gaps 0;



QY 1 ERKRNTIHEFKSAKTTLIKIDPALKIKTKKVNNTADQCANRCTRNKGLPFTCKAFVFDK 60  
Db :|||||  
QY 32 QKRRNTIHEFKSAKTTLIKIDPALKIKTKKVNNTADQCANRCTRNKGLPFTCKAFVFDK 91  
Db :|||||  
QY 61 ARKQCLWFPFNSMSSGVKKEGHEFDLYENKDYIRNCIIGKRSYKGTVSIKSGIKCOP 120  
Db :|||||  
QY 92 ARKQCLWFPFNSMSSGVKKEGHEFDLYENKDYIRNCIIGKRSYKGTVSIKSGIKCOP 151  
Db :|||||  
QY 121 WSSMIPHEHSFLPSSYRGKDLQENYCRNPRGEGGPGWCFSTSNPEVRYEVCIDIPOCSEVBC 180  
Db :|||||  
QY 152 WSSMIPHEHSFLPSSYRGKDLQENYCRNPRGEGGPGWCFSTSNPEVRYEVCIDIPOCSEVBC 211  
Db :|||||  
QY 181 MTCNGESYRGLMDHTESGKICQWHDQTPHRHKFLPERYPDKGFDNYCRNPDGQPRPMC 240  
Db :|||||  
QY 212 MTCNGESYRGLMDHTESGKICQWHDQTPHRHKFLPERYPDKGFDNYCRNPDGQPRPMC 271  
Db :|||||  
QY 241 YTLDPHTRWEXCAIKTCADNTMNDTDVPLETTECIQOGGEGYRGTVNTIWNIGPCQRWDS 300  
Db :|||||  
QY 272 YTLDPHTRWEXCAIKTCADNTMNDTDVPLETTECIQOGGEGYRGTVNTIWNIGPCQRWDS 331  
Db :|||||  
QY 301 QYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSOIPNCDMSHGQDC 360  
Db :|||||  
QY 332 QYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSOIPNCDMSHGQDC 391  
Db :|||||  
QY 361 YRNGKNTYMGNLQSOTRSGLTCSMDKNMEDLHRHIFWEPDASKLNEYCRNPDGDAHGPW 420  
Db :|||||  
QY 392 YRNGKNTYMGNLQSOTRSGLTCSMDKNMEDLHRHIFWEPDASKLNEYCRNPDGDAHGPW 451  
Db :|||||  
QY 421 CYTGNPLIPWDYCPISRCGEGDTTPTIV 447  
Db :|||||  
QY 452 CYTGNPLIPWDYCPISRCGEGDTTPTIV 478  
Db :|||||

Search completed: April 25, 2003, 15:50:29  
Job time : 83 secs

**THIS PAGE BLANK (uspto)**

GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 25, 2003, 15:40:22 ; Search time 18 seconds  
(without alignments)  
730.669 Million cell updates/sec

Title: US-09-674-377B-1  
Perfect score: 2504  
Sequence: 1 ERKRRNTIHEFKSAKTLI.....IPWDYCPISRCGDTTPTIV 447

Scoring table: BLOSUM62  
Gapop 10.0 ..Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued\_Patents\_AA.\*

- 1: /cgn2\_6/ptodata/1/iaa/5A.COMB.pep.\*
- 2: /cgn2\_6/ptodata/1/iaa/5B.COMB.pep.\*
- 3: /cgn2\_6/ptodata/1/iaa/6A.COMB.pep.\*
- 4: /cgn2\_6/ptodata/1/iaa/6B.COMB.pep.\*
- 5: /cgn2\_6/ptodata/1/iaa/PCTUS.COMB.pep.\*
- 6: /cgn2\_6/ptodata/1/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2601	99.9	697	2	US-08-460-890A-50
2	2601	99.9	697	3	US-08-167-641C-50
3	2601	99.9	697	4	US-08-460-971A-50
4	2601	99.9	697	4	US-08-462-040-50
5	2601	99.9	728	1	US-07-815-333A-2
6	2601	99.9	728	1	US-08-087-783A-22
7	2563.5	98.4	723	1	US-07-838-410-1
8	2563.5	98.4	723	1	US-08-290-937B-1
9	2563.5	98.4	723	1	US-08-290-937B-2
10	2563.5	98.4	723	1	US-08-404-643-1
11	2563.5	98.4	723	4	US-08-194-326-1
12	2560.5	98.3	723	4	US-08-030-410-3
13	2557.5	98.2	723	1	US-08-290-937B-3
14	2541.5	97.6	723	4	US-08-194-326-2
15	2541.5	97.6	723	4	US-08-700-519J-18
16	2539.5	97.5	723	4	US-08-194-326-3
17	2539.5	97.5	723	4	US-08-700-519J-19
18	1180.5	45.3	711	1	US-08-184-012C-8
19	1180.5	45.3	711	1	US-08-334-177-2
20	1180.5	45.3	711	2	US-08-666-082B-1
21	1180.5	45.3	711	2	US-08-766-982-2
22	1180.5	45.3	711	4	US-08-296-219-2
23	1180.5	45.3	711	5	PCT-US95-13830-2
24	1125	43.2	716	2	US-08-766-982-1
25	1125	43.2	716	4	US-09-296-219-1
26	1059	40.7	812	1	US-08-248-629A-1
27	1059	40.7	812	1	US-08-451-932-1

28	1059	40.7	812	1	US-08-452-260-1
29	1059	40.7	812	1	US-08-326-785-1
30	1059	40.7	812	2	US-08-612-788-1
31	1059	40.7	812	2	US-08-605-598B-1
32	1059	40.7	812	2	US-08-429-743-1
33	1059	40.7	812	2	US-08-866-735-1
34	1059	40.7	812	3	US-09-066-028-1
35	1059	40.7	812	5	PCT-US95-05107-1
36	1045.5	40.1	790	2	US-08-469-486-54
37	1045.5	40.1	790	2	US-08-469-658-54
38	1042	40.0	791	1	US-08-643-219-1
39	1042	40.0	791	2	US-09-131-995-1
40	1042	40.0	791	2	US-08-832-087B-1
41	1042	40.0	791	3	US-08-851-350-1
42	1042	40.0	791	4	US-09-132-154-1
43	1042	40.0	810	1	US-07-854-603-2
44	1042	40.0	810	1	US-08-147-000B-29
45	1042	40.0	810	4	US-09-086-514-1

#### ALIGNMENTS

RESULT 1  
US-08-460-890A-50  
; Sequence 50, Application US/08460890A  
; Patent No. 5994109  
; GENERAL INFORMATION:  
; APPLICANT: Woo, Savio L.C.  
; APPLICANT: Smith, Louis C.  
; APPLICANT: Cristiano, Richard J.  
; APPLICANT: Gottchalk, Stephen  
; TITLE OF INVENTION: NUCLEIC ACID TRANSPORTER SYSTEMS AND  
; TITLE OF INVENTION: METHODS OF USE  
; NUMBER OF SEQUENCES: 65  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Lyon & Lyon  
; STREET: 633 West Fifth Street  
; STREET: Suite 4700  
; CITY: Los Angeles  
; STATE: California  
; COUNTRY: U.S.A.  
; ZIP: 90071-2066  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
; MEDIUM TYPE: storage  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: IBM P.C. DOS 5.0  
; SOFTWARE: FastSeq for Windows 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/460.890A  
; FILING DATE: June 5, 1995  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/167,641  
; FILING DATE: December 14, 1993  
; APPLICATION NUMBER: 07/855,389  
; FILING DATE: March 20, 1992  
; APPLICATION NUMBER: PCT/US93/02725  
; FILING DATE: March 19, 1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Warburg, Richard J.  
; REGISTRATION NUMBER: 32,327  
; REFERENCE/DOCKET NUMBER: 212/066  
; TELEPHONE: (213) 489-1600  
; TELEFAX: (213) 955-0440  
; TELEX: 67-3510  
; INFORMATION FOR SEQ ID NO: 50:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 697 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single

```
;
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-460-830A-50

Query Match          99.9%; Score 2601; DB 2; Length 697;
Best Local Similarity 99.8%; Pred. No. 1.2e-214;
Matches 446; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ERKRRNTHEFKSAKTTLIKIDPALKIKTKKVTADOCANRCTRNKGLPTCKAFVFDK 60
DB 1 QKRRNTHEFKSAKTTLIKIDPALKIKTKKVTADOCANRCTRNKGLPTCKAFVFDK 60

QY 61 ARKQCLWFPFNSMSSGVKKEFGHEFDLYENKDYIRNCIIIGKRSYKGTVSITKSGIKQOP 120
DB 61 ARKQCLWFPFNSMSSGVKKEFGHEFDLYENKDYIRNCIIIGKRSYKGTVSITKSGIKQOP 120

QY 121 WSSMIPHEHSEFLPSSYRGKDLQENYCRNPRGEGGFWCFTSNPEVRYEVCIDIPOCSEVEC 180
DB 121 WSSMIPHEHSEFLPSSYRGKDLQENYCRNPRGEGGFWCFTSNPEVRYEVCIDIPOCSEVEC 180

QY 181 MTCNGESYRGLMDHTESGKICORWDHQPPIRHKFLPERYPDKGFDNCRNPDGQPRPWC 240
DB 181 MTCNGESYRGLMDHTESGKICORWDHQPPIRHKFLPERYPDKGFDNCRNPDGQPRPWC 240

QY 241 YTLDPHTRWEYCAIKTCADNTMNDTDVPLETTECIQGGEGYRGTVNTIWNIGIPQORWDS 300
DB 241 YTLDPHTRWEYCAIKTCADNTMNDTDVPLETTECIQGGEGYRGTVNTIWNIGIPQORWDS 300

QY 301 QYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDMSHGQDC 360
DB 301 QYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDMSHGQDC 360

QY 361 YRGNKKNYMGNLSTQTRSGLTCSMDKNMEDLHRHIFWEPDASKLKNENYCRNPDDDDAHGPW 420
DB 361 YRGNKKNYMGNLSTQTRSGLTCSMDKNMEDLHRHIFWEPDASKLKNENYCRNPDDDDAHGPW 420

QY 421 CYTGNPLIPWDYCPISRCBGDTTPTIV 447
DB 421 CYTGNPLIPWDYCPISRCBGDTTPTIV 447

RESULT 2
US-08-167-641C-50
; Sequence 50, Application US/08167641C
; Patent No. 6033884
; GENERAL INFORMATION:
; APPLICANT: Woo, Savio L.C.
; APPLICANT: Smith, Louis C.
; APPLICANT: Cristiano, Richard J.
; APPLICANT: Gottchalk, Stephen
; TITLE OF INVENTION: NUCLEIC ACID TRANSPORTER SYSTEMS AND
; NUMBER OF SEQUENCES: 65
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071-2066
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; MEDIUM TYPE: storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: FastSeq for Windows 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/167,641C
; FILING DATE: December 14, 1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/855,389

;
; FILING DATE: March 20, 1992
; APPLICATION NUMBER: PCT/US93/02725
; FILING DATE: March 19, 1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Warburg, Richard J.
; REGISTRATION NUMBER: 32,327
; REFERENCE/DOCKET NUMBER: 205/012
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 50:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 697 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-167-641C-50

Query Match          99.9%; Score 2601; DB 3; Length 697;
Best Local Similarity 99.8%; Pred. No. 1.2e-214;
Matches 446; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ERKRRNTHEFKSAKTTLIKIDPALKIKTKKVTADOCANRCTRNKGLPTCKAFVFDK 60
DB 1 QKRRNTHEFKSAKTTLIKIDPALKIKTKKVTADOCANRCTRNKGLPTCKAFVFDK 60

QY 61 ARKQCLWFPFNSMSSGVKKEFGHEFDLYENKDYIRNCIIIGKRSYKGTVSITKSGIKQOP 120
DB 61 ARKQCLWFPFNSMSSGVKKEFGHEFDLYENKDYIRNCIIIGKRSYKGTVSITKSGIKQOP 120

QY 121 WSSMIPHEHSEFLPSSYRGKDLQENYCRNPRGEGGFWCFTSNPEVRYEVCIDIPOCSEVEC 180
DB 121 WSSMIPHEHSEFLPSSYRGKDLQENYCRNPRGEGGFWCFTSNPEVRYEVCIDIPOCSEVEC 180

QY 181 MTCNGESYRGLMDHTESGKICORWDHQPPIRHKFLPERYPDKGFDNCRNPDGQPRPWC 240
DB 181 MTCNGESYRGLMDHTESGKICORWDHQPPIRHKFLPERYPDKGFDNCRNPDGQPRPWC 240

QY 241 YTLDPHTRWEYCAIKTCADNTMNDTDVPLETTECIQGGEGYRGTVNTIWNIGIPQORWDS 300
DB 241 YTLDPHTRWEYCAIKTCADNTMNDTDVPLETTECIQGGEGYRGTVNTIWNIGIPQORWDS 300

QY 301 QYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDMSHGQDC 360
DB 301 QYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDMSHGQDC 360

QY 361 YRGNKKNYMGNLSTQTRSGLTCSMDKNMEDLHRHIFWEPDASKLKNENYCRNPDDDDAHGPW 420
DB 361 YRGNKKNYMGNLSTQTRSGLTCSMDKNMEDLHRHIFWEPDASKLKNENYCRNPDDDDAHGPW 420

QY 421 CYTGNPLIPWDYCPISRCBGDTTPTIV 447
DB 421 CYTGNPLIPWDYCPISRCBGDTTPTIV 447

RESULT 3
US-08-460-971A-50
; Sequence 50, Application US/08460971A
; Patent No. 6150168
; GENERAL INFORMATION:
; APPLICANT: Woo, Savio L.C.
; APPLICANT: Smith, Louis C.
; APPLICANT: Cristiano, Richard J.
; APPLICANT: Gottchalk, Stephen
; TITLE OF INVENTION: NUCLEIC ACID TRANSPORTER SYSTEMS AND
; NUMBER OF SEQUENCES: 65
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071-2066
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; MEDIUM TYPE: storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: FastSeq for Windows 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/167,641C
; FILING DATE: December 14, 1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/855,389
```

```

: CITY: Los Angeles
: STATE: California
: COUNTRY: U.S.A.
: ZIP: 90071-2066
: COMPUTER READABLE FORM:
: MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
: MEDIUM TYPE: storage
: COMPUTER: IBM Compatible
: OPERATING SYSTEM: IBM P.C. DOS 5.0
: SOFTWARE: FastSEQ for Windows 2.0
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/460.971A
: FILING DATE: June 5, 1995
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 08/167.641
: FILING DATE: December 14, 1993
: APPLICATION NUMBER: 07/855.389
: FILING DATE: March 20, 1992
: APPLICATION NUMBER: PCT/US93/02725
: FILING DATE: March 19, 1993
: ATTORNEY/AGENT INFORMATION:
: NAME: Warburg, Richard J.
: REGISTRATION NUMBER: 32,327
: REFERENCE/DOCKET NUMBER: 212/063
: TELEPHONE: (213) 489-1600
: TELEFAX: (213) 955-0440
: TELEX: 67-3510
: INFORMATION FOR SEQ ID NO: 50:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 697 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: peptide
: US-08-460-971A-50

Query Match          99.9%; Score 2601; DB 4; Length 697;
Best Local Similarity 99.8%; Pred. No. 1.2e-214;
Matches 446; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ERKRNTTHEPKSAKTLIKIDPALKIKTKKVNADQCANCRNKGLPFTCKAFVFDK 60
DB :|||||
DB 1 QKRKNTTHEPKSAKTLIKIDPALKIKTKKVNADQCANCRNKGLPFTCKAFVFDK 60
QY 61 ARKQCLWPFNSMSSGVKKEFGHEFDLYENKDYIRNCIIIGKGRSVKGTVITKSGIKCOP 120
DB :|||||
DB 61 ARKQCLWPFNSMSSGVKKEFGHEFDLYENKDYIRNCIIIGKGRSVKGTVITKSGIKCOP 120
QY 121 WSSMIPHEHSFLPSSYRGKDLQENYCRNPRGEGGPGWCFTSNPEVRYEYCDIPQCSEVEC 180
DB :|||||
DB 121 WSSMIPHEHSFLPSSYRGKDLQENYCRNPRGEGGPGWCFTSNPEVRYEYCDIPQCSEVEC 180
QY 181 MTCNGESYRGLMDHTESGKICQWHDQTPHRRKFLPERYPDKGFDNFCRNPDGQPRPWC 240
DB :|||||
DB 181 MTCNGESYRGLMDHTESGKICQWHDQTPHRRKFLPERYPDKGFDNFCRNPDGQPRPWC 240
QY 241 YTLDPHTTWEYCAIKTCADNTMNDTVDPLETTECLOGGEGYRGVNTIWNIGIPCORWDS 300
DB :|||||
DB 241 YTLDPHTTWEYCAIKTCADNTMNDTVDPLETTECLOGGEGYRGVNTIWNIGIPCORWDS 300
QY 301 QYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDMSHGQDC 360
DB :|||||
DB 301 QYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDMSHGQDC 360
QY 361 YRGNGKNTYGNLSQTRSGLTCSMWKDNKEDLHRHIFWEPDASKLNEYCRNPDHAGHPW 420
DB :|||||
DB 361 YRGNGKNTYGNLSQTRSGLTCSMWKDNKEDLHRHIFWEPDASKLNEYCRNPDHAGHPW 420
QY 421 CYTGNPLIPWDYCPISRCGEGDTTPTIV 447
DB :|||||
DB 421 CYTGNPLIPWDYCPISRCGEGDTTPTIV 447

: CITY: Los Angeles
: STATE: California
: COUNTRY: U.S.A.
: ZIP: 90071-2066
: COMPUTER READABLE FORM:
: MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
: MEDIUM TYPE: storage
: COMPUTER: IBM Compatible
: OPERATING SYSTEM: IBM P.C. DOS 5.0
: SOFTWARE: FastSEQ for Windows 2.0
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/462.040
: FILING DATE: June 5, 1995
: CLASSIFICATION: 536
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 08/167.641
: FILING DATE: December 14, 1993
: APPLICATION NUMBER: 07/855.389
: FILING DATE: March 20, 1992
: APPLICATION NUMBER: PCT/US93/02725
: FILING DATE: March 19, 1993
: ATTORNEY/AGENT INFORMATION:
: NAME: Warburg, Richard J.
: REGISTRATION NUMBER: 32,327
: REFERENCE/DOCKET NUMBER: 212/078
: TELEPHONE: (213) 489-1600
: TELEFAX: (213) 955-0440
: TELEX: 67-3510
: INFORMATION FOR SEQ ID NO: 50:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 697 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: peptide
: US-08-462-040-50

Query Match          99.9%; Score 2601; DB 4; Length 697;
Best Local Similarity 99.8%; Pred. No. 1.2e-214;
Matches 446; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ERKRNTTHEPKSAKTLIKIDPALKIKTKKVNADQCANCRNKGLPFTCKAFVFDK 60
DB :|||||
DB 1 QKRKNTTHEPKSAKTLIKIDPALKIKTKKVNADQCANCRNKGLPFTCKAFVFDK 60
QY 61 ARKQCLWPFNSMSSGVKKEFGHEFDLYENKDYIRNCIIIGKGRSVKGTVITKSGIKCOP 120
DB :|||||
DB 61 ARKQCLWPFNSMSSGVKKEFGHEFDLYENKDYIRNCIIIGKGRSVKGTVITKSGIKCOP 120
QY 121 WSSMIPHEHSFLPSSYRGKDLQENYCRNPRGEGGPGWCFTSNPEVRYEYCDIPQCSEVEC 180
DB :|||||
DB 121 WSSMIPHEHSFLPSSYRGKDLQENYCRNPRGEGGPGWCFTSNPEVRYEYCDIPQCSEVEC 180
QY 181 MTCNGESYRGLMDHTESGKICQWHDQTPHRRKFLPERYPDKGFDNFCRNPDGQPRPWC 240
DB :|||||
DB 181 MTCNGESYRGLMDHTESGKICQWHDQTPHRRKFLPERYPDKGFDNFCRNPDGQPRPWC 240
QY 241 YTLDPHTTWEYCAIKTCADNTMNDTVDPLETTECLOGGEGYRGVNTIWNIGIPCORWDS 300
DB :|||||
DB 241 YTLDPHTTWEYCAIKTCADNTMNDTVDPLETTECLOGGEGYRGVNTIWNIGIPCORWDS 300
QY 301 QYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDMSHGQDC 360
DB :|||||
DB 301 QYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDMSHGQDC 360
QY 361 YRGNGKNTYGNLSQTRSGLTCSMWKDNKEDLHRHIFWEPDASKLNEYCRNPDHAGHPW 420
DB :|||||
DB 361 YRGNGKNTYGNLSQTRSGLTCSMWKDNKEDLHRHIFWEPDASKLNEYCRNPDHAGHPW 420
QY 421 CYTGNPLIPWDYCPISRCGEGDTTPTIV 447
DB :|||||
DB 421 CYTGNPLIPWDYCPISRCGEGDTTPTIV 447
```

121 WSSMIPHEHFLPSSYRGKDLQENYCRNPRGEGGWPCFTSNPEVRYEYCDIPQCSEVEC 180  
152 WSSMIPHEHFLPSSYRGKDLQENYCRNPRGEGGWPCFTSNPEVRYEYCDIPQCSEVEC 211  
181 MTNCGESYRGLMDHTESGKICQWDHQTTPHHRKFLPERYDPKGFDDNYCRNPDGQPRPWC 240  
212 MTNCGESYRGLMDHTESGKICQWDHQTTPHHRKFLPERYDPKGFDDNYCRNPDGQPRPWC 271  
241 YLDPHTRWEYCAIKTCADNTMDTVPLETTETECIOGEGYRGTVNTIWNIGPCQRWDS 300  
272 YLDPHTRWEYCAIKTCADNTMDTVPLETTETECIOGEGYRGTVNTIWNIGPCQRWDS 331  
301 QYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDMSHGQDC 360  
332 QYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDMSHGQDC 391  
361 YRGNKKNYMGNSLQTSRGLTCSMDKKNMEDLHRHIFWEPDASKLNENYCRNPDGDAHGPW 420  
392 YRGNKKNYMGNSLQTSRGLTCSMDKKNMEDLHRHIFWEPDASKLNENYCRNPDGDAHGPW 451  
421 CYTGNPLIPWDYCPISRCBGDTTPTIV 447  
452 CYTGNPLIPWDYCPISRCBGDTTPTIV 478

RESULT 5  
US-07-815-333A-2  
; Sequence 2, Application US/07815333A  
; Patent No. 5342831  
; GENERAL INFORMATION:  
; APPLICANT: Nakamura, Toshikazu  
; APPLICANT: Matsumoto, Kunio  
; TITLE OF INVENTION: EPITHELIOCYTE GROWTH ACCELERATOR  
; NUMBER OF SEQUENCES: 2  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Leydig, Voit & Mayer  
; STREET: Two Prudential Plaza, Suite 4900  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: United States of America  
; ZIP: 60601  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.25  
; CURRENT APPLICATION NUMBER: US/07/815,333A  
; FILING DATE: 19911227  
; CLASSIFICATION: 530  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Green, Robert F.  
; REGISTRATION NUMBER: 27555  
; REFERENCE/DOCKET NUMBER: 44069  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (312) 616-5600  
; TELEFAX: (312) 616-5700  
; TELEX: 25-3533  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 728 amino acids  
; TYPE: AMINO ACID  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-07-815-333A-2

Query Match 99.9%; Score 2601; DB 1; Length 728;  
Best Local Similarity 99.8%; Pred. No. 1.2e-214;  
Matches 446; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ERKRRNTIHEPKSAKTTLIKIDPALIKTKKVNATDOCANRCTRNKGLPFTCKAFVFDK 60  
DB 32 QKRRNTIHEPKSAKTTLIKIDPALIKTKKVNATDOCANRCTRNKGLPFTCKAFVFDK 91

QY 61 ARKQCLWFPNFMSSGVKKEGHEFDLYENKDYIRNCIIIGKRSYKGTVTSITSGIKCQP 120  
DB 92 ARKQCLWFPNFMSSGVKKEGHEFDLYENKDYIRNCIIIGKRSYKGTVTSITSGIKCQP 151

121 WSSMIPHEHFLPSSYRGKDLQENYCRNPRGEGGWPCFTSNPEVRYEYCDIPQCSEVEC 180  
152 WSSMIPHEHFLPSSYRGKDLQENYCRNPRGEGGWPCFTSNPEVRYEYCDIPQCSEVEC 211  
181 MTNCGESYRGLMDHTESGKICQWDHQTTPHHRKFLPERYDPKGFDDNYCRNPDGQPRPWC 240  
212 MTNCGESYRGLMDHTESGKICQWDHQTTPHHRKFLPERYDPKGFDDNYCRNPDGQPRPWC 271  
241 YLDPHTRWEYCAIKTCADNTMDTVPLETTETECIOGEGYRGTVNTIWNIGPCQRWDS 300  
272 YLDPHTRWEYCAIKTCADNTMDTVPLETTETECIOGEGYRGTVNTIWNIGPCQRWDS 331  
301 QYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDMSHGQDC 360  
332 QYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDMSHGQDC 391  
361 YRGNKKNYMGNSLQTSRGLTCSMDKKNMEDLHRHIFWEPDASKLNENYCRNPDGDAHGPW 420  
392 YRGNKKNYMGNSLQTSRGLTCSMDKKNMEDLHRHIFWEPDASKLNENYCRNPDGDAHGPW 451  
421 CYTGNPLIPWDYCPISRCBGDTTPTIV 447  
452 CYTGNPLIPWDYCPISRCBGDTTPTIV 478

RESULT 6  
US-08-087-783A-22  
; Sequence 22, Application US/08087783A  
; Patent No. 5547856  
; GENERAL INFORMATION:  
; APPLICANT: Godowski, Paul J., Lokker, Nathalie A., Mark, Melanie R.  
; TITLE OF INVENTION: HEPATOCYTE GROWTH FACTOR VARIANTS  
; NUMBER OF SEQUENCES: 22  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Genentech, Inc.  
; STREET: 460 Point San Bruno Blvd  
; CITY: South San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94080  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 3.5 inch, 1.44 mb floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Winpatin (Genentech)  
; CURRENT APPLICATION NUMBER: US/08/087,783A  
; FILING DATE: 13-Jul-1993  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/884811  
; FILING DATE: 18-MAY-92  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/885971  
; FILING DATE: 18-MAY-92  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Marschang, Diane L.  
; REGISTRATION NUMBER: 35,600  
; REFERENCE/DOCKET NUMBER: P0755779P1  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415/225-5416  
; TELEFAX: 415/952-9881  
; TELEX: 910/371-7168  
; INFORMATION FOR SEQ ID NO: 22:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 728 amino acids  
; TYPE: Amino Acid  
; TOPOLOGY: Linear  
US-08-087-783A-22

Query Match 99.9%; Score 2601; DB 1; Length 728;  
Best Local Similarity 99.8%; Pred. No. 1.2e-214;  
Matches 446; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ERKRRNTIHEFKSAKTTLIKIDPALKIKTKKVTADQCANCTRNKGLPFTCKAFVFDK 60  
:|||||  
Db 32 QRKRNTIHEFKSAKTTLIKIDPALKIKTKKVTADQCANCTRNKGLPFTCKAFVFDK 91  
:|||||  
QY 61 ARKQCLWFPFNSMSGVKKFEGHEFDLYENKDYIRNCIIGKRSYKGTYSITKSGIKCOP 120  
:|||||  
Db 92 ARKQCLWFPFNSMSGVKKFEGHEFDLYENKDYIRNCIIGKRSYKGTYSITKSGIKCOP 151  
:|||||  
QY 121 WSSMIPHEHSFLPSSYRGKDLQENCRNPRGEGGPGWCFSTSNPEVRYEYCDIPQCSEVEC 180  
:|||||  
Db 152 WSSMIPHEHSFLPSSYRGKDLQENCRNPRGEGGPGWCFSTSNPEVRYEYCDIPQCSEVEC 211  
:|||||  
QY 181 MTCNGESYRGLMDHTESKICORWDHQTTPHRRKFLPERYPDKGFDNCRNPDGQPRPWC 240  
:|||||  
Db 212 MTCNGESYRGLMDHTESKICORWDHQTTPHRRKFLPERYPDKGFDNCRNPDGQPRPWC 271  
:|||||  
QY 241 YTLDPHTRWEYCAIKTCADNTMNDTVDPLETTECIQGGEGYRGVTNTIWNIGIPCORWDS 300  
:|||||  
Db 272 YTLDPHTRWEYCAIKTCADNTMNDTVDPLETTECIQGGEGYRGVTNTIWNIGIPCORWDS 331  
:|||||  
QY 301 QYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDMSHGQDC 360  
:|||||  
Db 332 QYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDMSHGQDC 391  
:|||||  
QY 361 YRGNGKNYMGNLQSOTRSGLTCSMWKDNMEDLHRHIFWEPDASKNLNENCRNPDGDAHGFW 420  
:|||||  
Db 392 YRGNGKNYMGNLQSOTRSGLTCSMWKDNMEDLHRHIFWEPDASKNLNENCRNPDGDAHGFW 451  
:|||||  
QY 421 CYTGNPLIPWDYCPISRCBGDTTPTIV 447  
:|||||  
Db 452 CYTGNPLIPWDYCPISRCBGDTTPTIV 478  
:|||||

RESULT 7  
US-07-838-410-1  
: Sequence 1, Application US/07838410  
: Patent No. 5328836  
: GENERAL INFORMATION:  
: APPLICANT: SHIMA, NOBUYUKI  
: APPLICANT: HIGASHIO, KANJI  
: APPLICANT: NAGAO, MASAYA  
: APPLICANT: OOGAKI, FUMIKO  
: APPLICANT: TAKAOKA, HIROAKI  
: APPLICANT: TSUDA, ELSUKE  
: TITLE OF INVENTION: PLASMIDS CONTAINING DNA ENCODING THE  
: TITLE OF INVENTION: AMINO ACID SEQUENCE OF TCF-II, TRANSFORMED CELLS WITH THE  
: TITLE OF INVENTION: PLASMIDS, AND PRODUCTION METHOD OF BIOLOGICALLY ACTIVE  
: TITLE OF INVENTION: SUBSTANCE USING THE TRANSFORMED CELLS  
: NUMBER OF SEQUENCES: 11  
: CORRESPONDENCE ADDRESS:  
: ADDRESSEE: MILLEN, WHITE, ZELANO & BRANIGAN, P.C.  
: STREET: 2200 CLARENDON BLVD.  
: CITY: ARLINGTON  
: STATE: VA  
: COUNTRY: USA  
: ZIP: 22201  
: COMPUTER READABLE FORM:  
: MEDIUM TYPE: Floppy disk  
: COMPUTER: IBM PC compatible  
: OPERATING SYSTEM: PC-DOS/MS-DOS  
: SOFTWARE: PatentIn Release #1.0, Version #1.25  
: CURRENT APPLICATION DATA:  
: APPLICATION NUMBER: US/07/838.410  
: FILING DATE: 19920311  
: CLASSIFICATION: 435  
: PRIOR APPLICATION DATA:  
: APPLICATION NUMBER: PCT/JP91/00942  
: FILING DATE: 15-JUL-1991  
: PRIOR APPLICATION DATA:  
: APPLICATION NUMBER: JP 185852-1990  
: FILING DATE: 13-JUL-1990  
: ATTORNEY/AGENT INFORMATION:

: NAME: WHITE, JOHN L.  
: REGISTRATION NUMBER: 17,746  
: REFERENCE/DOCKET NUMBER: WAK 110  
: TELECOMMUNICATION INFORMATION:  
: TELEPHONE: 703-243-6333  
: TELEFAX: 703-243-6410  
: TELEX: 64191  
: INFORMATION FOR SEQ ID NO: 1:  
: SEQUENCE CHARACTERISTICS:  
: LENGTH: 723 amino acids  
: TYPE: AMINO ACID  
: STRANDEDNESS: single  
: TOPOLOGY: linear  
: MOLECULE TYPE: protein  
: ORIGINAL SOURCE:  
: ORGANISM: Homo sapiens  
: CELL TYPE: Fibroblast  
: FEATURE:  
: NAME/KEY: Domain  
: LOCATION: 393..405  
: OTHER INFORMATION: /note= "INTERNAL AMINO ACID  
: OTHER INFORMATION: SEQUENCE IN ALPHA-CHAIN"  
: FEATURE:  
: NAME/KEY: Protein  
: LOCATION: 490..505  
: OTHER INFORMATION: /note= "N-TERMINAL AMINO ACID  
: OTHER INFORMATION: SEQUENCE OF BETA-CHAIN"  
: FEATURE:  
: NAME/KEY: Domain  
: LOCATION: 605..623  
: OTHER INFORMATION: /note= "INTERNAL AMINO ACID  
: OTHER INFORMATION: SEQUENCE IN BETA-CHAIN"  
: US-07-838-410-1

Query Match 98.4%; Score 2563.5; DB 1; Length 723;  
Best Local Similarity 98.7%; Pred No. 2e-211; Mismatches 1; Gaps 1;  
Matches 441; Conservative 1; Indels 5; Gaps 1;  
QY 1 ERKRRNTIHEFKSAKTTLIKIDPALKIKTKKVTADQCANCTRNKGLPFTCKAFVFDK 60  
:|||||  
Db 32 QRKRNTIHEFKSAKTTLIKIDPALKIKTKKVTADQCANCTRNKGLPFTCKAFVFDK 91  
:|||||  
QY 61 ARKQCLWFPFNSMSGVKKFEGHEFDLYENKDYIRNCIIGKRSYKGTYSITKSGIKCOP 120  
:|||||  
Db 92 ARKQCLWFPFNSMSGVKKFEGHEFDLYENKDYIRNCIIGKRSYKGTYSITKSGIKCOP 151  
:|||||  
QY 121 WSSMIPHEHSFLPSSYRGKDLQENCRNPRGEGGPGWCFSTSNPEVRYEYCDIPQCSEVEC 180  
:|||||  
Db 152 WSSMIPHEH-----SYRGKDLQENCRNPRGEGGPGWCFSTSNPEVRYEYCDIPQCSEVEC 206  
:|||||  
QY 181 MTCNGESYRGLMDHTESKICORWDHQTTPHRRKFLPERYPDKGFDNCRNPDGQPRPWC 240  
:|||||  
Db 207 MTCNGESYRGLMDHTESKICORWDHQTTPHRRKFLPERYPDKGFDNCRNPDGQPRPWC 266  
:|||||  
QY 241 YTLDPHTRWEYCAIKTCADNTMNDTVDPLETTECIQGGEGYRGVTNTIWNIGIPCORWDS 300  
:|||||  
Db 267 YTLDPHTRWEYCAIKTCADNTMNDTVDPLETTECIQGGEGYRGVTNTIWNIGIPCORWDS 326  
:|||||  
QY 301 QYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDMSHGQDC 360  
:|||||  
Db 327 QYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDMSHGQDC 386  
:|||||  
QY 361 YRGNGKNYMGNLQSOTRSGLTCSMWKDNMEDLHRHIFWEPDASKNLNENCRNPDGDAHGFW 420  
:|||||  
Db 387 YRGNGKNYMGNLQSOTRSGLTCSMWKDNMEDLHRHIFWEPDASKNLNENCRNPDGDAHGFW 446  
:|||||  
QY 421 CYTGNPLIPWDYCPISRCBGDTTPTIV 447  
:|||||  
Db 447 CYTGNPLIPWDYCPISRCBGDTTPTIV 473  
:|||||

RESULT 8  
US-08-290-937B-1

Sequence 1, Application US/08290937B	Sequence 1, Application US/08290937B-2
Patent No. 5648233	Patent No. 5648233
GENERAL INFORMATION:	GENERAL INFORMATION:
APPLICANT: YAMAGUCHI, KYOJI	APPLICANT: YAMAGUCHI, KYOJI
APPLICANT: SHIMA, NOBUYUKI	APPLICANT: SHIMA, NOBUYUKI
APPLICANT: MURAKAMI, AKIHIKO	APPLICANT: MURAKAMI, AKIHIKO
APPLICANT: GOTO, MASAOKI	APPLICANT: GOTO, MASAOKI
APPLICANT: TSUDA, EISUKE	APPLICANT: TSUDA, EISUKE
APPLICANT: MASUNAGA, HIROAKI	APPLICANT: MASUNAGA, HIROAKI
APPLICANT: TAKAHIRA, REIKO	APPLICANT: TAKAHIRA, REIKO
APPLICANT: OOGAKI, FUMIKO	APPLICANT: OOGAKI, FUMIKO
APPLICANT: UEDA, MASATSUGU	APPLICANT: UEDA, MASATSUGU
APPLICANT: HIGASHIO, KANJI	APPLICANT: HIGASHIO, KANJI
TITLE OF INVENTION: MODIFIED TCF	TITLE OF INVENTION: MODIFIED TCF
NUMBER OF SEQUENCES: 13	NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:	CORRESPONDENCE ADDRESS:
ADDRESSEE: Testa, Hurwitz & Thibault	ADDRESSEE: Testa, Hurwitz & Thibault
STREET: 125 High St.	STREET: 125 High St.
CITY: Boston	CITY: Boston
STATE: MA	STATE: MA
COUNTRY: USA	COUNTRY: USA
ZIP: 02110	ZIP: 02110
COMPUTER READABLE FORM:	COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk	MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible	COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS	OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30	SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:	CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/290, 937B	APPLICATION NUMBER: US/08/290, 937B
FILING DATE: 19-AUG-1994	FILING DATE: 19-AUG-1994
CLASSIFICATION: 530	CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:	ATTORNEY/AGENT INFORMATION:
NAME: CAMPBELL, PAULA A.	NAME: CAMPBELL, PAULA A.
REGISTRATION NUMBER: 32,503	REGISTRATION NUMBER: 32,503
REFERENCE/DOCKET NUMBER: FJN-022	REFERENCE/DOCKET NUMBER: FJN-022
TELECOMMUNICATION INFORMATION:	TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 248-7000	TELEPHONE: (617) 248-7000
TELEFAX: (617) 248-7100	TELEFAX: (617) 248-7100
INFORMATION FOR SEQ ID NO: 1:	INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:	SEQUENCE CHARACTERISTICS:
LENGTH: 723 amino acids	LENGTH: 723 amino acids
TYPE: amino acid	TYPE: amino acid
STRANDEDNESS: single	STRANDEDNESS: single
TOPOLOGY: linear	TOPOLOGY: linear
MOLECULE TYPE: protein	MOLECULE TYPE: protein
US-08-290-937B-1	US-08-290-937B-2
Query Match 98.4%; Score 2563.5; DB 1; Length 723;	Query Match 98.4%; Score 2563.5; DB 1; Length 723;
Best Local Similarity 98.7%; Pred. No. 2e-211;	Best Local Similarity 98.7%; Pred. No. 2e-211;
Matches 441; Conservative 1; Mismatches 0; Indels 5; Gaps 1;	Matches 441; Conservative 1; Mismatches 0; Indels 5; Gaps 1;
QY 1 ERKRRNTIHEFKKSAKTTLIKIDPALKIKTKKVVNTADOCANCRTRNKGLPTCKAFVFDK 60	QY 1 ERKRRNTIHEFKKSAKTTLIKIDPALKIKTKKVVNTADOCANCRTRNKGLPTCKAFVFDK 60
DB 32 QRKRRNTIHEFKKSAKTTLIKIDPALKIKTKKVVNTADOCANCRTRNKGLPTCKAFVFDK 91	DB 32 QRKRRNTIHEFKKSAKTTLIKIDPALKIKTKKVVNTADOCANCRTRNKGLPTCKAFVFDK 91
QY 61 ARKQCLWFFNSMSGVKKFEGHEFDLYENKDYIRNCLIGKGRSYKGTIVTSITKSGIKCP 120	QY 61 ARKQCLWFFNSMSGVKKFEGHEFDLYENKDYIRNCLIGKGRSYKGTIVTSITKSGIKCP 120
DB 92 ARKQCLWFFNSMSGVKKFEGHEFDLYENKDYIRNCLIGKGRSYKGTIVTSITKSGIKCP 151	DB 92 ARKQCLWFFNSMSGVKKFEGHEFDLYENKDYIRNCLIGKGRSYKGTIVTSITKSGIKCP 151
QY 121 WSSMIPHEHSFLPSSYRGKDLQENYCRNPRGEGGPGWFTSNPEVRYEVCDDIPOCSEVEC 180	QY 121 WSSMIPHEHSFLPSSYRGKDLQENYCRNPRGEGGPGWFTSNPEVRYEVCDDIPOCSEVEC 180
DB 152 WSSMIPHEH-----SYRGKDLQENYCRNPRGEGGPGWFTSNPEVRYEVCDDIPOCSEVEC 206	DB 152 WSSMIPHEH-----SYRGKDLQENYCRNPRGEGGPGWFTSNPEVRYEVCDDIPOCSEVEC 206
QY 181 MTCNGESYRGLMDHTESGICORWDHQTPHRRKFLPERYPDKGDDNNYCRNPDGQPRWC 240	QY 181 MTCNGESYRGLMDHTESGICORWDHQTPHRRKFLPERYPDKGDDNNYCRNPDGQPRWC 240
DB 207 MTCNGESYRGLMDHTESGICORWDHQTPHRRKFLPERYPDKGDDNNYCRNPDGQPRWC 266	DB 207 MTCNGESYRGLMDHTESGICORWDHQTPHRRKFLPERYPDKGDDNNYCRNPDGQPRWC 266
QY 241 YTLDPHTRWEYCAIKTCADNTWNTDVPLETTTCOGGEGYRGVNTVWNGICQWRDS 300	QY 241 YTLDPHTRWEYCAIKTCADNTWNTDVPLETTTCOGGEGYRGVNTVWNGICQWRDS 300
DB 267 YTLDPHTRWEYCAIKTCADNTWNTDVPLETTTCOGGEGYRGVNTVWNGICQWRDS 326	DB 267 YTLDPHTRWEYCAIKTCADNTWNTDVPLETTTCOGGEGYRGVNTVWNGICQWRDS 326
QY 301 QYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDMSHGDC 360	QY 301 QYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDMSHGDC 360



Db 152 WSSMIPHEH-----SYRGKDLQENYCRNPRGEGGPGWCFSTSNPEVRYEYCDIPQCSEVEC 206  
QY 181 MTCNGESYRGLMDHTESGKICORWDHQTTPHRRKFLPERYPDKGFDNDYCRNPDGQPRPWC 240  
Db 207 MTCNGESYRGLMDHTESGKICORWDHQTTPHRRKFLPERYPDKGFDNDYCRNPDGQPRPWC 266  
QY 241 YTLDPHTRWYCAIKTCADNTMNDTVDVPLETTECIGQGEGYRGVTNTIWNIGIPQCRWDS 300  
Db 267 YTLDPHTRWYCAIKTCADNTMNDTVDVPLETTECIGQGEGYRGVTNTIWNIGIPQCRWDS 326  
QY 301 QYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDMSHGQDC 360  
Db 327 QYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDMSHGQDC 386  
QY 361 YRGNGKNTMGNLSQTRSGLTCSMWKDNMEDLHRHIFWEPDASKLNEYCRNPDGQPRPWC 420  
Db 387 YRGNGKNTMGNLSQTRSGLTCSMWKDNMEDLHRHIFWEPDASKLNEYCRNPDGQPRPWC 446  
QY 421 CYTGNPLIPWDYCPISRCBGDTTPTIV 447  
Db 447 CYTGNPLIPWDYCPISRCBGDTTPTIV 473

## RESULT 10

US-08-404-643-1  
; Sequence 1, Application US/08404643  
; Patent No. 5658742  
; GENERAL INFORMATION:  
; APPLICANT: HIGASHIO, KANJI  
; APPLICANT: SHIMA, NOBUYUKI  
; APPLICANT: OOGAKI, FUMIKO  
; TITLE OF INVENTION: MONOCLONAL ANTIBODY  
; NUMBER OF SEQUENCES: 1  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: PATENT ADMINISTRATOR, TESTA, HURWITZ &  
; ADDRESSEE: THIBEAULT  
; STREET: 53 STATE STREET  
; CITY: BOSTON  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02109

COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/404,643  
; FILING DATE: 15-MAR-1995  
; CLASSIFICATION: 436  
; ATTORNEY/AGENT INFORMATION:  
; NAME: CAMPBELL, PAULA A  
; REGISTRATION NUMBER: 32,503  
; REFERENCE/DOCKET NUMBER: FJN-034 (3999/35)  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617)-248-7000  
; TELEFAX: (617)-248-7100  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 723 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-404-643-1

.Query Match 98.4%; Score 2563.5; DB 1; Length 723;  
Best Local Similarity 98.7%; Pred. No. 2e-211;  
Matches 441; Conservative 1; Mismatches 0; Indels 5; Gaps 1;  
QY 1 ERKRNTTHEFKKSAKTTLIKIDPALKIKTKKVNADQCANRCTRNKGLPFTCKAFVFDK 60  
:|||||

Db 32 QKRRTTHEFKKSAKTTLIKIDPALKIKTKKVNADQCANRCTRNKGLPFTCKAFVFDK 91  
QY 61 ARKQCLWFPFNSMSSGVKKEFGHEFDLYENKDYIRNCIIGKGRSYKGTYSITKSGIKCOP 120  
Db 92 ARKQCLWFPFNSMSSGVKKEFGHEFDLYENKDYIRNCIIGKGRSYKGTYSITKSGIKCOP 151  
QY 121 WSSMIPHEHSFLPSSYRGKDLQENYCRNPRGEGGPGWCFSTSNPEVRYEYCDIPQCSEVEC 180  
Db 152 WSSMIPHEH-----SYRGKDLQENYCRNPRGEGGPGWCFSTSNPEVRYEYCDIPQCSEVEC 206  
QY 181 MTCNGESYRGLMDHTESGKICORWDHQTTPHRRKFLPERYPDKGFDNDYCRNPDGQPRPWC 240  
Db 207 MTCNGESYRGLMDHTESGKICORWDHQTTPHRRKFLPERYPDKGFDNDYCRNPDGQPRPWC 266  
QY 241 YTLDPHTRWYCAIKTCADNTMNDTVDVPLETTECIGQGEGYRGVTNTIWNIGIPQCRWDS 300  
Db 267 YTLDPHTRWYCAIKTCADNTMNDTVDVPLETTECIGQGEGYRGVTNTIWNIGIPQCRWDS 326  
QY 301 QYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDMSHGQDC 360  
Db 327 QYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDMSHGQDC 386  
QY 361 YRGNGKNTMGNLSQTRSGLTCSMWKDNMEDLHRHIFWEPDASKLNEYCRNPDGQPRPWC 420  
Db 387 YRGNGKNTMGNLSQTRSGLTCSMWKDNMEDLHRHIFWEPDASKLNEYCRNPDGQPRPWC 446  
QY 421 CYTGNPLIPWDYCPISRCBGDTTPTIV 447  
Db 447 CYTGNPLIPWDYCPISRCBGDTTPTIV 473

## RESULT 11

US-09-194-326-1  
; Sequence 1, Application US/09194326  
; Patent No. 6306827  
; GENERAL INFORMATION:  
; APPLICANT: Kinoshita, Masahiko  
; APPLICANT: Ogawa, Hiroshi  
; APPLICANT: Masanaga, Hiroaki  
; APPLICANT: Kobayashi, Fumie  
; APPLICANT: Yamaguchi, Kyoji  
; APPLICANT: Higashio, Kanji  
; TITLE OF INVENTION: Agent for Preventing and/or Treating Renal Disease  
; FILE REFERENCE: FJN-069  
; CURRENT APPLICATION NUMBER: US/09/194,326  
; CURRENT FILING DATE: 1998-11-24  
; EARLIER APPLICATION NUMBER: PCT/JP98/01221  
; EARLIER FILING DATE: 1998-03-20  
; EARLIER APPLICATION NUMBER: JP 94989  
; EARLIER FILING DATE: 1997-03-28  
; NUMBER OF SEQ ID NOS: 3  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 1  
; LENGTH: 723  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: TCF-II  
US-09-194-326-1

.Query Match 98.4%; Score 2563.5; DB 4; Length 723;  
Best Local Similarity 98.7%; Pred. No. 2e-211;  
Matches 441; Conservative 1; Mismatches 0; Indels 5; Gaps 1;

QY 1 ERKRNTTHEFKKSAKTTLIKIDPALKIKTKKVNADQCANRCTRNKGLPFTCKAFVFDK 60  
:|||||  
Db 32 QKRRTTHEFKKSAKTTLIKIDPALKIKTKKVNADQCANRCTRNKGLPFTCKAFVFDK 91  
QY 61 ARKQCLWFPFNSMSSGVKKEFGHEFDLYENKDYIRNCIIGKGRSYKGTYSITKSGIKCOP 120  
Db 92 ARKQCLWFPFNSMSSGVKKEFGHEFDLYENKDYIRNCIIGKGRSYKGTYSITKSGIKCOP 151  
QY 121 WSSMIPHEHSFLPSSYRGKDLQENYCRNPRGEGGPGWCFSTSNPEVRYEYCDIPQCSEVEC 180

		Matches, 440:	Conservative	2:	Mismatches	0:	Indels	5:	Gaps	1:
Db	152	WSSMIPHEH-----SYRGKDLQENYCRNPRGEGPWCFTSNPEVRYEVCDIPOCSEVEC	206							
Qy	181	MTCNGESYRGLMDHTESGKICQWRDQHPHRRKFLPERYPDKGDDNDYCRNPDGQPRPWC	240							
Db	207	MTCNGESYRGLMDHTESGKICQWRDQHPHRRKFLPERYPDKGDDNDYCRNPDGQPRPWC	266							
Qy	241	YTLDPHTRWEYCAIKTCADNTMNDTVPLETTTCIOGOGEGYRGTVNTIWNIGIPCQRWDS	300							
Db	267	YTLDPHTRWEYCAIKTCADNTMNDTVPLETTTCIOGOGEGYRGTVNTIWNIGIPCQRWDS	326							
Qy	301	QYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDMSHGQDC	360							
Db	327	QYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDMSHGQDC	386							
Qy	361	YRGNKGYMGNLSQTRSGLTCSMWDKNMEDLHRRHIFWEPDASKLKNENYCRNPDHDAHPW	420							
Db	387	YRGNKGYMGNLSQTRSGLTCSMWDKNMEDLHRRHIFWEPDASKLKNENYCRNPDHDAHPW	446							
Qy	421	CYTGNPLIPWDYCPISRCGEGDTTPTIV	447							
Db	447	CYTGNPLIPWDYCPISRCGEGDTTPTIV	473							

```
RESULT 13
US-08-290-937B-3
; Sequence 3, Application US/08290937B
; Patent No. 5648233
; GENERAL INFORMATION:
; APPLICANT: YAMAGUCHI, KYOJI
; APPLICANT: SHIMA, NOBUYUKI
; APPLICANT: MURAKAMI, AKIHIKO
; APPLICANT: GOTO, MASAOKI
; APPLICANT: TSUDA, EISUKE
; APPLICANT: MASUNAGA, HIROAKI
; APPLICANT: TAKAHIRA, REIKO
; APPLICANT: OOGAKI, FUMIKO
; APPLICANT: UEDA, MASATSUGU
; APPLICANT: HIGASHIO, KANJI
; TITLE OF INVENTION: MODIFIED TCF
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Testa, Hurwitz & Thibault
; STREET: 125 High St.
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/290,937B
; FILING DATE: 19-AUG-1994
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: CAMPBELL, PAULA A.
; REGISTRATION NUMBER: 32,503
; REFERENCE/DOCKET NUMBER: FJN-022
; TELECOMMUNICATION INFORMATION:
```

```

; TELEPHONE: (617) 248-7000
; TELEFAX: (617) 248-7100
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 723 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-290-937B-3

Query Match          98.2%; Score 2557.5; DB 1; Length 723;
Best Local Similarity 98.4%; Pred. No. 6.5e-211;
Matches 440; Conservative 1; Mismatches 1; Indels 5; Gaps 1;

QY 1 ERKRNTIHEFKSAKTTLIKIDPALKIKTKKVTADQCANRCTRNKGLPFTCKAFVFDK 60
DB 32 QRKRNTIHEFKSAKTTLIKIDPALKIKTKKVTADQCANRCTRNKGLPFTCKAFVFDK 91
QY 61 ARKQCLWFPNMSGGVKKEFGHEFDLYENKDYIRNCIIIGKGRSYKGTVSTTKSGIKQCP 120
DB 92 ARKQCLWFPNMSGGVKKEFGHEFDLYENKDYIRNCIIIGKGRSYKGTVSTTKSGIKQCP 151
QY 121 WSSMIPHEHSLPSSYRGKDLQENYCRNPRGEGGWCFTSNPEVRYEVCIDIPOCSEVEC 180
DB 152 WSSMIPHEH-----SYRGKDLQENYCRNPRGEGGWCFTSNPEVRYEVCIDIPOCSEVEC 206
QY 181 MTCNGESYRGLMDHTESGKICQWDHQTPHRHKFLPERYPDKGFDNNYCRNPDGQPRPWC 240
DB 207 MTCNGESYRGLMDHTESGKICQWDHQTPHRHKFLPERYPDKGFDNNYCRNPDGQPRPWC 266
QY 241 YLDPHTRWEYCAIKTCADNTMNDTVPLETTETCIGQGGEGYRGTVNTIWNIGPCQWRDS 300
DB 267 YLDPHTRWEYCAIKTCADNTMNDTVPLETTETCIGQGGEGYRGTVNTIWNIGPCQWRDS 326
QY 301 QYPHEHMTPENFCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDMSHGQDC 360
DB 327 QYPHEHMTPENFCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDMSHGQDC 386
QY 361 YRGNKYNMGNLSQTRSGLTCSMWDKNMEDLHRHIFWEPDASKLNENYCRNPDHAGPWCYTGN 420
DB 387 YRGNKYNMGNLSQTRSGLTCSMWDKNMEDLHRHIFWEPDASKLNENYCRNPDHAGPWCYTGN 456
QY 421 CYTGNPLIPWDYCPISRCGEGDTTPTIV 447
DB 447 CYTGNPLIPWDYCPISRCGEGDTTPTIV 473

RESULT 14
US-09-194-326-2
; Sequence 2, Application US/09194326
; Patent No. 6306827
; GENERAL INFORMATION:
; APPLICANT: Kinoshaki, Masahiko
; APPLICANT: Ogawa, Hiromi
; APPLICANT: Masanaga, Hiroaki
; APPLICANT: Kobayashi, Fumie
; APPLICANT: Yamaguchi, Kyoji
; APPLICANT: Higashio, Kanji
; TITLE OF INVENTION: Agent for Preventing and/or Treating Renal Disease
; FILE REFERENCE: FJN-069
; CURRENT APPLICATION NUMBER: US/09/194,326
; CURRENT FILING DATE: 1998-11-24
; EARLIER APPLICATION NUMBER: PCT/JF98/01221
; EARLIER FILING DATE: 1998-03-20
; EARLIER APPLICATION NUMBER: JP 94989
; EARLIER FILING DATE: 1997-03-28
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn ver. 2.0
; SEQ ID NO 2
; LENGTH: 723
; TYPE: PRT
; ORGANISM: Artificial Sequence

; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: RKRR2AAAA
; OTHER INFORMATION: mutant of TCF-II
US-09-194-326-2

Query Match          97.68; Score 2541.5; DB 4; Length 723;
Best Local Similarity 98.9%; Pred. No. 1.5e-209;
Matches 437; Conservative 0; Mismatches 0; Indels 5; Gaps 1;

QY 6 NTIHEFKSAKTTLIKIDPALKIKTKKVTADQCANRCTRNKGLPFTCKAFVFDKARKQC 65
DB 37 NTIHEFKSAKTTLIKIDPALKIKTKKVTADQCANRCTRNKGLPFTCKAFVFDKARKQC 96
QY 66 LWFPNMSGGVKKEFGHEFDLYENKDYIRNCIIIGKGRSYKGTVSTTKSGIKQCPWSSMI 125
DB 97 LWFPNMSGGVKKEFGHEFDLYENKDYIRNCIIIGKGRSYKGTVSTTKSGIKQCPWSSMI 156
QY 126 PHEHSLPSSYRGKDLQENYCRNPRGEGGWCFTSNPEVRYEVCIDIPOCSEVECMTGNG 185
DB 157 PHEH-----SYRGKDLQENYCRNPRGEGGWCFTSNPEVRYEVCIDIPOCSEVECMTGNG 211
QY 186 ESYRGLMDHTESGKICQWDHQTPHRHKFLPERYPDKGFDNNYCRNPDGQPRPWCYTLDLP 245
DB 212 ESYRGLMDHTESGKICQWDHQTPHRHKFLPERYPDKGFDNNYCRNPDGQPRPWCYTLDLP 271
QY 246 HTRWEYCAIKTCADNTMNDTVPLETTETCIGQGGEGYRGTVNTIWNIGPCQWRDSQYPHE 305
DB 272 HTRWEYCAIKTCADNTMNDTVPLETTETCIGQGGEGYRGTVNTIWNIGPCQWRDSQYPHE 331
QY 306 HDMTPENFCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDMSHGQDCYRNG 365
DB 332 HDMTPENFCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDMSHGQDCYRNG 391
QY 366 KNYMGNLSQTRSGLTCSMWDKNMEDLHRHIFWEPDASKLNENYCRNPDHAGPWCYTGN 425
DB 392 KNYMGNLSQTRSGLTCSMWDKNMEDLHRHIFWEPDASKLNENYCRNPDHAGPWCYTGN 451
QY 426 PLIPWDYCPISRCGEGDTTPTIV 447
DB 452 PLIPWDYCPISRCGEGDTTPTIV 473

RESULT 15
US-08-700-519J-18
; Sequence 18, Application US/08700519J
; Patent No. 6395744
; GENERAL INFORMATION:
; APPLICANT: Kinoshaki, Masahiko
; APPLICANT: Yamaguchi, Kyoji
; APPLICANT: Goto, Masaaki
; APPLICANT: Murakami, Akihiko
; APPLICANT: Ueda, Masatsugu
; APPLICANT: Higashio, Kanji
; APPLICANT: Yamashita, Yasushi
; TITLE OF INVENTION: TCF MUTANT
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESS: Burgess, Ryan and Wayne
; STREET: 370 Lexington Avenue, Suite 2105
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 10017
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3 1/5 inch diskette
; COMPUTER: PC'S LIMITED SYSTEM 200
; OPERATING SYSTEM: WORD FOR WINDOWS
; SOFTWARE: MICROSOFT WINDOWS 98
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/700,519J
; FILING DATE: 26-Aug-1996
; CLASSIFICATION: <unknown>
; PRIOR APPLICATION DATA:

```

Search completed: April 25, 2003, 15:45:23  
Job time : 23 secs

GenCore version 5.1.4.p5.4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 25, 2003, 15:44:38 ; Search time 21 seconds  
(without alignments)  
1705.620 Million cell updates/sec

Title: US-09-674-377B-1

Perfect score: 2604

Sequence: 1 ERKRNTIHEFKSAKTLI.....IPWDYCPISRCGDTTPTIV 447

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 301932 seqs, 80129803 residues

Total number of hits satisfying chosen parameters: 301932

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications\_AA:\*

- 1: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB pep.\*
- 2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB pep.\*
- 3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB pep.\*
- 4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB pep.\*
- 5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB pep.\*
- 6: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB pep.\*
- 7: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB pep.\*
- 8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB pep.\*
- 9: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB pep.\*
- 10: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB pep.\*
- 11: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB pep.\*
- 12: /cgn2\_6/ptodata/1/pubpaa/US10\_PUBCOMB pep.\*
- 13: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB pep.\*
- 14: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	2601	99.9	447	9	US-10-081-309-2
2	2601	99.9	728	1	US-08-605-221-2
3	2563.5	98.4	723	1	US-08-605-221-4
4	2541.5	97.6	723	9	US-10-133-912-18
5	2539.5	97.5	723	9	US-10-133-912-19
6	1180.5	45.3	711	9	US-10-133-036-2
7	1059	40.7	812	9	US-09-335-325-1
8	1059	40.7	812	9	US-10-131-241-1
9	1059	40.7	812	10	US-09-788-142-1
10	1059	40.7	812	10	US-09-761-120-1
11	1059	40.7	812	10	US-09-873-676-81
12	1042	40.0	569	10	US-09-946-893-5
13	1042	40.0	571	10	US-09-946-893-8
14	1042	40.0	576	10	US-09-946-893-6
15	1042	40.0	791	9	US-09-967-386-1
16	1042	40.0	810	10	US-09-946-893-2
17	1034	39.7	451	9	US-10-157-369-2
18	1022.5	39.3	378	9	US-09-335-325-41
19	1022.5	39.3	378	9	US-10-131-241-41

20	1022.5	39.3	378	10	US-09-761-120-41	Sequence 41, Appl
21	1003.5	38.5	459	10	US-09-761-120-46	Sequence 46, Appl
22	1003	38.5	352	9	US-09-335-325-39	Sequence 39, Appl
23	1003	38.5	352	9	US-10-131-241-39	Sequence 39, Appl
24	1003	38.5	352	10	US-09-761-120-39	Sequence 39, Appl
25	988.5	38.0	378	9	US-09-335-325-42	Sequence 42, Appl
26	988.5	38.0	378	9	US-10-131-241-42	Sequence 42, Appl
27	988.5	38.0	378	10	US-09-873-676-1	Sequence 1, Appl
28	985	37.8	364	9	US-10-157-369-4	Sequence 4, Appl
29	966.5	37.1	458	10	US-09-946-893-4	Sequence 4, Appl
30	962.5	37.0	368	10	US-09-761-120-42	Sequence 42, Appl
31	960	36.9	352	9	US-09-335-325-40	Sequence 40, Appl
32	960	36.9	352	9	US-10-131-241-40	Sequence 40, Appl
33	960	36.9	352	10	US-09-761-120-40	Sequence 40, Appl
34	942	36.2	339	9	US-09-335-325-2	Sequence 2, Appl
35	942	36.2	339	9	US-10-131-241-2	Sequence 2, Appl
36	942	36.2	339	10	US-09-788-142-2	Sequence 2, Appl
37	942	36.2	339	10	US-09-761-120-2	Sequence 2, Appl
38	940	36.1	339	9	US-09-335-325-5	Sequence 5, Appl
39	940	36.1	339	9	US-10-131-241-5	Sequence 5, Appl
40	940	36.1	339	10	US-09-788-142-5	Sequence 5, Appl
41	940	36.1	339	10	US-09-761-120-5	Sequence 5, Appl
42	929	35.7	423	12	US-10-036-869-23	Sequence 23, Appl
43	922	35.4	339	9	US-09-335-325-6	Sequence 6, Appl
44	922	35.4	339	9	US-10-131-241-6	Sequence 6, Appl
45	922	35.4	339	10	US-09-788-142-6	Sequence 6, Appl

#### ALIGNMENTS

##### RESULT 1

US-10-081-309-2

; Sequence 2, Application US/10081309

; Publication No. US20030012775A1

; GENERAL INFORMATION:

; APPLICANT: Hoffmann-La Roche Inc.

; TITLE OF INVENTION: PEG Conjugates of NK4

; FILE REFERENCE: 20859

; CURRENT APPLICATION NUMBER: US/10/081.309

; CURRENT FILING DATE: 2002-06-04

; NUMBER OF SEQ ID NOS: 2

; SOFTWARE: Patentin version 3.1

; SEQ ID NO 2

; LENGTH: 447

; TYPE: PRT

; ORGANISM: homomapiens

US-10-081-309-2

Query Match 99.9%; Score 2601; , DB 9; Length 447;

Best Local Similarity 99.8%; Pred. No. 6.4e-202;

Matches 446; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy	1	ERKRNTIHEFKSAKTLI	KIDPALKIKTKKVNADQCANR	CTRNKGLPFTCKAFVDK	60
Db	1	QRKRNTIHEFKSAKTLI	KIDPALKIKTKKVNADQCANR	CTRNKGLPFTCKAFVDK	60
Qy	61	ARKQCLWFPFNSMSGV	KKEFGHEFDLYENKDY	IRNCIIIGKGRSVKGT	VSITKSGIKQCP 120
Db	61	ARKQCLWFPFNSMSGV	KKEFGHEFDLYENKDY	IRNCIIIGKGRSVKGT	VSITKSGIKQCP 120
Qy	121	WSSMPIHESFLPSSY	RKGLQENYCRNPRGEGG	PWCFTSNPEVRYE	VCIDIPOCSEVEC 180
Db	121	WSSMPIHESFLPSSY	RKGLQENYCRNPRGEGG	PWCFTSNPEVRYE	VCIDIPOCSEVEC 180
Qy	181	MTNCGESYRGLMDH	TESGKICQWRDQHPH	RHKFLPERYPDKG	FDNDCRNPDGQRPWC 240
Db	181	MTNCGESYRGLMDH	TESGKICQWRDQHPH	RHKFLPERYPDKG	FDNDCRNPDGQRPWC 240
Qy	241	YTLDPHTRWEYCAI	KTCADNTMNDTVDV	PLETTECLOGQGE	YRGVTNTWNGIPCORWDS 300
Db	241	YTLDPHTRWEYCAI	KTCADNTMNDTVDV	PLETTECLOGQGE	YRGVTNTWNGIPCORWDS 300

```
QY 301 QYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDMSHGQDC 360
Db 301 QYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDMSHGQDC 360
QY 361 YRGNGKNYMGNSQTRSGLTCSMDKKNMEDLHRHIFWEPDASKLNNENYCRNPDDDAHGWP 420
Db 361 YRGNGKNYMGNSQTRSGLTCSMDKKNMEDLHRHIFWEPDASKLNNENYCRNPDDDAHGWP 420
QY 421 CYTGNPLIPWDYCPISRCBGDTTPTIV 447
Db 421 CYTGNPLIPWDYCPISRCBGDTTPTIV 447

RESULT 2
US-08-605-221-2
; Sequence 2, Application US/08605221
; Publication No. US20030060403A1
; GENERAL INFORMATION:
; APPLICANT: NAKAMURA, Toshikazu
; TITLE OF INVENTION: THERAPEUTIC AGENT FOR DISORDER IN BRAIN AND NERVE
; FILE REFERENCE: 2520-0101P
; CURRENT APPLICATION NUMBER: US/08/605,221
; CURRENT FILING DATE: 1996-03-15
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 728
; TYPE: PRT
; ORGANISM: Homo sapiens
; OTHER INFORMATION: HGF-A (Full length HGF)
US-08-605-221-2

Query Match 99.9%; Score 2601; DB 1; Length 728;
Best Local Similarity 99.8%; Pred. No. 1.2e-201;
Matches 446; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ERKRNTIHEFKSAKTLTIKIDPALKIKTKKVNADQCANRCTRNKGLPFTCKAFVFDK 60
Db 32 QKRKRNTIHEFKSAKTLTIKIDPALKIKTKKVNADQCANRCTRNKGLPFTCKAFVFDK 91
QY 61 ARKQCLWFPFNSMSGYKKEFGHEFDLYENKDYIRNCILIGKRSYKGTVSTKSGIKQCP 120
Db 92 ARKQCLWFPFNSMSGYKKEFGHEFDLYENKDYIRNCILIGKRSYKGTVSTKSGIKQCP 151
QY 121 WSSMIPHEHSLPSSYRGKDLQENYCRNPRGEGGWCFTSNPEVRYEVCDDIPQCSEVEC 180
Db 152 WSSMIPHEHSLPSSYRGKDLQENYCRNPRGEGGWCFTSNPEVRYEVCDDIPQCSEVEC 206
QY 181 MTCNGESYRGLMDHTESGKICQRDWDHQPTRHKFLPERYPDKGFDNNYCRNPDGQPRPWC 240
Db 207 MTCNGESYRGLMDHTESGKICQRDWDHQPTRHKFLPERYPDKGFDNNYCRNPDGQPRPWC 266
QY 241 YLDPHTRWECYCAIKTCADNTMNDTVPLETTTCIOGEGYRGTVNTIWNIGIPCORWDS 300
Db 267 YLDPHTRWECYCAIKTCADNTMNDTVPLETTTCIOGEGYRGTVNTIWNIGIPCORWDS 326
QY 301 QYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDMSHGQDC 360
Db 327 QYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDMSHGQDC 386
QY 361 YRGNGKNYMGNSQTRSGLTCSMDKKNMEDLHRHIFWEPDASKLNNENYCRNPDDDAHGWP 420
Db 387 YRGNGKNYMGNSQTRSGLTCSMDKKNMEDLHRHIFWEPDASKLNNENYCRNPDDDAHGWP 446
QY 421 CYTGNPLIPWDYCPISRCBGDTTPTIV 447
Db 447 CYTGNPLIPWDYCPISRCBGDTTPTIV 473

RESULT 4
US-10-133-912-18
; Sequence 18, Application US/10133912
; Patent No. US20020165358A1
; GENERAL INFORMATION:
; APPLICANT: Kinoshita, Masahiko
; Yamaguchi, Kyoji
; Goto, Masaaki
; Murakami, Akihiko
; Ueda, Masatsugu
; Higashio, Kanji
; Yamashita, Yasushi
; TITLE OF INVENTION: TCF MUTANT
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESS: Burgess, Ryan and Wayne
; STREET: 370 Lexington Avenue, Suite 2105
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 10017
```

```
QY 301 QYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDMSHGQDC 360
Db 301 QYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDMSHGQDC 360
QY 361 YRGNGKNYMGNSQTRSGLTCSMDKKNMEDLHRHIFWEPDASKLNNENYCRNPDDDAHGWP 420
Db 361 YRGNGKNYMGNSQTRSGLTCSMDKKNMEDLHRHIFWEPDASKLNNENYCRNPDDDAHGWP 420
QY 421 CYTGNPLIPWDYCPISRCBGDTTPTIV 447
Db 421 CYTGNPLIPWDYCPISRCBGDTTPTIV 447

RESULT 2
US-08-605-221-4
; Sequence 2, Application US/08605221
; Publication No. US20030060403A1
; GENERAL INFORMATION:
; APPLICANT: NAKAMURA, Toshikazu
; TITLE OF INVENTION: THERAPEUTIC AGENT FOR DISORDER IN BRAIN AND NERVE
; FILE REFERENCE: 2520-0101P
; CURRENT APPLICATION NUMBER: US/08/605,221
; CURRENT FILING DATE: 1996-03-15
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 723
; TYPE: PRT
; ORGANISM: Homo sapiens
; OTHER INFORMATION: HGF-B (five amino acids deletion type)
US-08-605-221-4

Query Match 98.4%; Score 2563.5; DB 1; Length 723;
Best Local Similarity 98.7%; Pred. No. 1.2e-198;
Matches 441; Conservative 1; Mismatches 0; Indels 5; Gaps 1;

QY 1 ERKRNTIHEFKSAKTLTIKIDPALKIKTKKVNADQCANRCTRNKGLPFTCKAFVFDK 60
Db 32 QKRKRNTIHEFKSAKTLTIKIDPALKIKTKKVNADQCANRCTRNKGLPFTCKAFVFDK 91
QY 61 ARKQCLWFPFNSMSGYKKEFGHEFDLYENKDYIRNCILIGKRSYKGTVSTKSGIKQCP 120
Db 92 ARKQCLWFPFNSMSGYKKEFGHEFDLYENKDYIRNCILIGKRSYKGTVSTKSGIKQCP 151
QY 121 WSSMIPHEHSLPSSYRGKDLQENYCRNPRGEGGWCFTSNPEVRYEVCDDIPQCSEVEC 180
Db 152 WSSMIPHEHSLPSSYRGKDLQENYCRNPRGEGGWCFTSNPEVRYEVCDDIPQCSEVEC 206
QY 181 MTCNGESYRGLMDHTESGKICQRDWDHQPTRHKFLPERYPDKGFDNNYCRNPDGQPRPWC 240
Db 207 MTCNGESYRGLMDHTESGKICQRDWDHQPTRHKFLPERYPDKGFDNNYCRNPDGQPRPWC 266
QY 241 YLDPHTRWECYCAIKTCADNTMNDTVPLETTTCIOGEGYRGTVNTIWNIGIPCORWDS 300
Db 267 YLDPHTRWECYCAIKTCADNTMNDTVPLETTTCIOGEGYRGTVNTIWNIGIPCORWDS 326
QY 301 QYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDMSHGQDC 360
Db 327 QYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDMSHGQDC 386
QY 361 YRGNGKNYMGNSQTRSGLTCSMDKKNMEDLHRHIFWEPDASKLNNENYCRNPDDDAHGWP 420
Db 387 YRGNGKNYMGNSQTRSGLTCSMDKKNMEDLHRHIFWEPDASKLNNENYCRNPDDDAHGWP 446
QY 421 CYTGNPLIPWDYCPISRCBGDTTPTIV 447
Db 447 CYTGNPLIPWDYCPISRCBGDTTPTIV 473

RESULT 4
US-10-133-912-18
; Sequence 18, Application US/10133912
; Patent No. US20020165358A1
; GENERAL INFORMATION:
; APPLICANT: Kinoshita, Masahiko
; Yamaguchi, Kyoji
; Goto, Masaaki
; Murakami, Akihiko
; Ueda, Masatsugu
; Higashio, Kanji
; Yamashita, Yasushi
; TITLE OF INVENTION: TCF MUTANT
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESS: Burgess, Ryan and Wayne
; STREET: 370 Lexington Avenue, Suite 2105
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 10017
```

COMPUTER READABLE FORM:  
MEDIUM TYPE: 3 1/5 inch diskette  
COMPUTER: PC'S LIMITED SYSTEM 200  
OPERATING SYSTEM: WORD FOR WINDOWS  
SOFTWARE: MICROSOFT WINDOWS 98  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/133,912  
FILING DATE: 25-Apr-2002  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/700,519  
FILING DATE: 26-Aug-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Wayne, Milton J.  
REFERENCE/DOCKET NUMBER: U-Wp-5240PCT  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-683-8150  
TELEFAX: 212-532-4285  
INFORMATION FOR SEQ ID NO: 18:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 723  
TYPE: AMINO ACID  
STRANDEDNESS: SINGLE  
TOPOLOGY: LINEAR  
SEQUENCE DESCRIPTION: SEQ ID NO: 18:  
US-10-133-912-18

Query Match 97.6%; Score 2541.5; DB 9; Length 723;  
Best Local Similarity 98.9%; Pred. No. 7.2e-197;  
Matches 437; Conservative 0; Mismatches 0; Indels 5; Gaps 1;  
QY 6 NTIHEFKSAKTLIKIDPALKIKTKVNTADOCANRCKTRNKGLPFTCKAFVDFDKARQC 65  
Db 37 NTIHEFKSAKTLIKIDPALKIKTKVNTADOCANRCKTRNKGLPFTCKAFVDFDKARQC 96  
QY 66 LWPFNSMSSGVKKEFGHEFDLYENKDYIRNCIIIGKRSYKGTVSITKSGIKQPWSSMI 125  
Db 97 LWPFNSMSSGVKKEFGHEFDLYENKDYIRNCIIIGKRSYKGTVSITKSGIKQPWSSMI 156  
QY 126 PHHSFLPSSYRGKDLQENYCRNPRGEGGWCFTSNPEVRYEVCIDIPOCSEVECMTCNG 185  
Db 157 PHEH-----SYRGKDLQENYCRNPRGEGGWCFTSNPEVRYEVCIDIPOCSEVECMTCNG 211  
QY 186 ESYRGLMDHTESGKICQWHDQTPHRRKFLPERYPDKGFDNDYCRNPDGQPRPWCYTLDP 245  
Db 212 ESYRGLMDHTESGKICQWHDQTPHRRKFLPERYPDKGFDNDYCRNPDGQPRPWCYTLDP 271  
QY 246 HTRWEYCAITCADNTMNDTDVPLETTECICQGGEGYRGTVNTIWNIGIPCORWDSQYPHE 305  
Db 272 HTRWEYCAITCADNTMNDTDVPLETTECICQGGEGYRGTVNTIWNIGIPCORWDSQYPHE 331  
QY 306 HDMPENFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIIPNCDSHGDQCYRGNG 365  
Db 332 HDMPENFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIIPNCDSHGDQCYRGNG 391  
QY 366 KNYMGNLSQTRSGLTCSMWKDNMEDLHRHIFWEPDASKLNYENYCRNPDHAGHPWCYGTGN 425  
Db 392 KNYMGNLSQTRSGLTCSMWKDNMEDLHRHIFWEPDASKLNYENYCRNPDHAGHPWCYGTGN 451  
QY 426 PLIPWDYCPISRCGGDTTPTIV 447  
Db 452 PLIPWDYCPISRCGGDTTPTIV 473

RESULT 5  
US-10-133-912-19  
Sequence 19, Application US/10133912  
Patent No. US20020165358A1  
GENERAL INFORMATION:  
APPLICANT: Kinosaki, Masahiko  
Yamaguchi, Kyoji  
Goto, Masaaki

Murakami, Akihiko  
Ueda, Masatsugu  
Higashio, Kanji  
Yamashita, Yasushi  
TITLE OF INVENTION: TCF MUTANT  
NUMBER OF SEQUENCES: 25  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Burgess, Ryan and Wayne  
STREET: 370 Lexington Avenue, Suite 2105  
CITY: NEW YORK  
STATE: NEW YORK  
COUNTRY: UNITED STATES OF AMERICA  
ZIP: 10017  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3 1/5 inch diskette  
COMPUTER: PC'S LIMITED SYSTEM 200  
OPERATING SYSTEM: WORD FOR WINDOWS  
SOFTWARE: MICROSOFT WINDOWS 98  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/133,912  
FILING DATE: 25-Apr-2002  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/700,519  
FILING DATE: 26-Aug-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Wayne, Milton J.  
REGISTRATION NUMBER: 17,906  
REFERENCE/DOCKET NUMBER: U-Wp-5240PCT  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-683-8150  
TELEFAX: 212-532-4285  
INFORMATION FOR SEQ ID NO: 19:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 723  
TYPE: AMINO ACID  
STRANDEDNESS: SINGLE  
TOPOLOGY: LINEAR  
SEQUENCE DESCRIPTION: SEQ ID NO: 19:  
US-10-133-912-19

Query Match 97.5%; Score 2539.5; DB 9; Length 723;  
Best Local Similarity 97.8%; Pred. No. 1e-196;  
Matches 437; Conservative 1; Mismatches 4; Indels 5; Gaps 1;  
QY 1 ERKRNTIHEFKSAKTLIKIDPALKIKTKVNTADOCANRCKTRNKGLPFTCKAFVDFK 60  
Db 32 ERKRNTIHEFKSAKTLIKIDPALKIKTKVNTADOCANRCKTRNKGLPFTCKAFVDFK 91  
QY 61 ARKQCLWFPFNSMSSGVKKEFGHEFDLYENKDYIRNCIIIGKRSYKGTVSITKSGIKCQP 120  
Db 92 ARKQCLWFPFNSMSSGVKKEFGHEFDLYENKDYIRNCIIIGKRSYKGTVSITKSGIKCQP 151  
QY 121 WSMIPHEHFLPSSYRGKDLQENYCRNPRGEGGWCFTSNPEVRYEVCIDIPOCSEVEEC 180  
Db 152 WSMIPHEH-----SYRGKDLQENYCRNPRGEGGWCFTSNPEVRYEVCIDIPOCSEVEEC 206  
QY 181 MTCNGESYRGLMDHTESGKICQWHDQTPHRRKFLPERYPDKGFDNDYCRNPDGQPRPWC 240  
Db 207 MTCNGESYRGLMDHTESGKICQWHDQTPHRRKFLPERYPDKGFDNDYCRNPDGQPRPWC 266  
QY 241 YTLDPHTRWEYCAITCADNTMNDTDVPLETTECICQGGEGYRGTVNTIWNIGIPCORWDS 300  
Db 267 YTLDPHTRWEYCAITCADNTMNDTDVPLETTECICQGGEGYRGTVNTIWNIGIPCORWDS 326  
QY 301 QYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIIPNCDSHGDQ 360  
Db 327 QYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIIPNCDSHGDQ 386  
QY 361 YRGNKKNYMGNLSQTRSGLTCSMWKDNMEDLHRHIFWEPDASKLNYENYCRNPDHAGHPW 420  
Db 387 YRGNKKNYMGNLSQTRSGLTCSMWKDNMEDLHRHIFWEPDASKLNYENYCRNPDHAGHPW 446





Db 443 SVRWEYCNLKR-SETGGSVV 462

RESULT 8

US-10-131-241-1

Sequence 1, Application US/10131241

Publication No. US20030012792A1

GENERAL INFORMATION:

APPLICANT: Holaday, John W.

APPLICANT: Fortier, Anne H.

TITLE OF INVENTION: Compositions and Methods for Inhibiting Endothelial Cell Proliferation

FILE REFERENCE: 05213-0344 43170-271565

CURRENT APPLICATION NUMBER: US/10131241

CURRENT FILING DATE: 2002-07-22

PRIOR APPLICATION NUMBER: US 09/413,049

PRIOR FILING DATE: 1999-10-06

PRIOR APPLICATION NUMBER: US 09/316,802

PRIOR FILING DATE: 1999-05-21

PRIOR APPLICATION NUMBER: US 60/086,586

PRIOR FILING DATE: 1998-05-22

NUMBER OF SEQ ID NOS: 65

SOFTWARE: PatentIn version 3.1

SEQ ID NO 1

LENGTH: 812

TYPE: PRT

ORGANISM: Murinae sp.

US-10-131-241-1

Query Match 40.7%; Score 1059; DB 9; Length 812;

Best Local Similarity 43.3%; Pred. No. 2.8e-77;

Matches 191; Conservative 67; Mismatches 145; Indels 38; Gaps 12;

Qy 30 TKKVNTA--DQCANRCTRNKGLPFTCKAFVFDKARKOCLWFPFNSMSGVKKKEGHEFD 86

Db 37 TKQLAAGVSDCLAKCEG--TDFVCRSFQYHSKEQOCVMAENSKTSSIR--RDVI 92

Qy 87 LYENKDYIRNCIIIGKRGYKGTVSTITKGIKQCPWSSMIPHEHSLPSSYRGKDLQENYC 146

Db 93 LFEKRVYLSCKTGIGNGYRGTMSTRTSGVACQKKGATFPHVPNTSPSTHPNEGLEENYC 152

Qy 147 RNPREGGPGCFSTNPEVRYEVCDDIPOCSEVECTMGESYRGMLDMHTESGKICQWDH 206

Db 153 RNPNDDEGPGWCYITDPDKRYDYNICEE--ECMYCSEGEYEGKISKTMSGLDQAWDS 211

Qy 207 QTPHRHKLPERYPDKGFDNNCRNPDGQPRPWCYTLDPHTRWEYCAIKTCADNTMNDTD 266

Db 212 QSPHAGYIPAKFSPKLNKMYCHNPDEPRPWCFTTDPTRWEYCDIPRCT---TTPP 267

Qy 267 VPLETECQOGEGYRGTVNTIWNIGIPCQRWDSQYPHEHDMTPENFKCKDLRENYCRNP 326

Db 268 PPSPTYQCLKRGENYRGTVSVTSGKTCQRMSEQTPHRRNRTPENFPCKNLEENYC 327

Qy 327 DGSSESPWCFTTDPNIRVGYCSQIPNCDMSHG-----ODCYRGNGKNYM 369

Db 328 DGETAPWCYITDQSLRWEYC-EIPSCSSASPDQSDSSVPEEQTPVVOEQYQSDGQSYR 386

Qy 370 GNLSQTRSLGCSMWDKNNEDLHRHIFWE---PDASKLNENYCRNPDDAHGPGWCYTCNP 426

Db 387 GTSSTITGKKCQSWAAMP--HRHSKTPENFPDAG-LEMNYCRNPDDG-KGPWCYITDP 442

Qy 427 LIPNDYCPISCEGDTTPIV 447

Db 443 SVRWEYCNLKR-SETGGSVV 462

RESULT 9

US-09-788-142-1

Sequence 1, Application US/09788142

Patent No. US20010029246A1

GENERAL INFORMATION:

APPLICANT: Folkman, M. Judah

O'Reilly, Michael

TITLE OF INVENTION: Angiostatin Fragments and Method of Use

NUMBER OF SEQUENCES: 6

CORRESPONDENCE ADDRESS:

ADDRESSEE: Jones & Askew, LLP

STREET: 191 Peachtree Street, 37th Floor

CITY: Atlanta

STATE: Georgia

COUNTRY: USA

ZIP: 30303-1769

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/788,142

FILING DATE: 16-Feb-2001

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/866,735

FILING DATE: <Unknown>

ATTORNEY/AGENT INFORMATION:

NAME: Warren, William L.

REGISTRATION NUMBER: 36,714

REFERENCE/DOCKET NUMBER: 05940-0129

TELECOMMUNICATION INFORMATION:

TELEPHONE: (404) 818-3700

TELEFAX: (404) 818-3799

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 812 amino acids

TYPE: amino acid

STRANDEDNESS: <Unknown>

TOPOLOGY: linear

MOLECULE TYPE: protein

HYPOTHETICAL: NO

ANTI-SENSE: NO

FRAGMENT TYPE: N-terminal

ORIGINAL SOURCE:

ORGANISM: Murine

IMMEDIATE SOURCE:

CLONE: Plasmidogen

SEQUENCE DESCRIPTION: SEQ ID NO: 1:

US-09-788-142-1

Query Match 40.7%; Score 1059; DB 10; Length 812;

Best Local Similarity 43.3%; Pred. No. 2.8e-77;

Matches 191; Conservative 67; Mismatches 145; Indels 38; Gaps 12;

Qy 30 TKKVNTA--DQCANRCTRNKGLPFTCKAFVFDKARKOCLWFPFNSMSGVKKKEGHEFD 86

Db 37 TKQLAAGVSDCLAKCEG--TDFVCRSFQYHSKEQOCVMAENSKTSSIR--RDVI 92

Qy 87 LYENKDYIRNCIIIGKRGYKGTVSTITKGIKQCPWSSMIPHEHSLPSSYRGKDLQENYC 146

Db 93 LFEKRVYLSCKTGIGNGYRGTMSTRTSGVACQKKGATFPHVPNTSPSTHPNEGLEENYC 152

Qy 147 RNPREGGPGCFSTNPEVRYEVCDDIPOCSEVECTMGESYRGMLDMHTESGKICQWDH 206

Db 153 RNPNDDEGPGWCYITDPDKRYDYNICEE--ECMYCSEGEYEGKISKTMSGLDQAWDS 211

Qy 207 QTPHRHKLPERYPDKGFDNNCRNPDGQPRPWCYTLDPHTRWEYCAIKTCADNTMNDTD 266

Db 212 QSPHAGYIPAKFSPKLNKMYCHNPDEPRPWCFTTDPTRWEYCDIPRCT---TTPP 267

Qy 267 VPLETECQOGEGYRGTVNTIWNIGIPCQRWDSQYPHEHDMTPENFKCKDLRENYCRNP 326

Db 268 PPSPTYQCLKRGENYRGTVSVTSGKTCQRMSEQTPHRRNRTPENFPCKNLEENYC 327

Qy 327 DGSSESPWCFTTDPNIRVGYCSQIPNCDMSHG-----ODCYRGNGKNYM 369

Db 328 DGETAPWCYITDQSLRWEYC-EIPSCSSASPDQSDSSVPEEQTPVVOEQYQSDGQSYR 386

QY	370	GNLSQTSRGLTCSMWKDNMEDLHRHFWB---	PDASKLNENYCRNPDDDDAGHPWCYTGNP	426
Db	387	GTSTTTITTKKCSWAAMPF--	HRHSKTPENFPDAG-LENNYCRNPDDG-KGPWCYTIDP	442
QY	427	LIPWDYCPISRCBGDTTPTIV	447	
Db	443	SVRWYECNLKRC-SETGGSVV	462	

```

RESULT 10
US-09-761-120-1
; Sequence 1, Application US/09761120
; Patent No. US20020037847A1
; GENERAL INFORMATION:
; APPLICANT: Folkman, M. Judah
; APPLICANT: O'Reilly, Michael
; TITLE OF INVENTION: Nucleic Acids Encoding Kringle 1-5 Region Fragments of Plasminogen
; FILE REFERENCE: 05940-0151 (43171-252068)
; CURRENT APPLICATION NUMBER: US/09/761,120
; CURRENT FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: 09/309,821
; PRIOR FILING DATE: 1999-05-11
; PRIOR APPLICATION NUMBER: 08/866,735
; PRIOR FILING DATE: 1997-05-30
; NUMBER OF SEQ ID NOS: 47
; - SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 812
; TYPE: PRT
; ORGANISM: Murinae gen. sp.
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Plasminogen
US-09-761-120-1

```

```

US-09-873-676-81
; Sequence 81, Application US/09873676
; Patent No. US20020077289A1
; GENERAL INFORMATION:
; APPLICANT: MacDonald, Nicholas J.
; APPLICANT: Sim, Kim L.
; TITLE OF INVENTION: Angiotensin and Endostatin Binding Proteins and Methods of Use
; FILE REFERENCE: 05213-0378 (43170-259333)
; CURRENT APPLICATION NUMBER: US/09/873,676
; CURRENT FILING DATE: 2001-06-04
; PRIOR APPLICATION NUMBER: US 60/209,065
; PRIOR FILING DATE: 2000-06-02
; PRIOR APPLICATION NUMBER: US 60/289,387
; PRIOR FILING DATE: 2001-05-08
; NUMBER OF SEQ ID NOS: 123
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 81
; LENGTH: 812
; TYPE: PRT
; ORGANISM: Murinae gen. sp.
US-09-873-676-81

```

RESULT 12  
US-09-946-893-5  
; Sequence 5, Application US/09946893  
; Patent No. US20020072494A1  
; GENERAL INFORMATION:  
; APPLICANT: Cao, Yi-hai  
; TITLE OF INVENTION: Materials and methods relating to endothelial cell growth  
; TITLE OF INVENTION: inhibitors  
; FILE REFERENCE: Newburn  
; CURRENT APPLICATION NUMBER: US/09/946, 893  
; CURRENT FILING DATE: 2001-09-05  
; PRIOR APPLICATION NUMBER: US 60/230, 893  
; PRIOR FILING DATE: 2000-09-05  
; NUMBER OF SEQ ID NOS: 11

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 5

; LENGTH: 569

; TYPE: PRT

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Description of Artificial Sequence: Angioquiescin  
US-09-946-893-5

Query Match 40.0%; Score 1042; DB 10; Length 569;

Best Local Similarity 43.6%; Pred. No. 4.3e-76;

Matches 193; Conservative 68; Mismatches 140; Indels 42; Gaps 13;

Qy 30 TKK---VNTADQCANRTRNKGLPFTCKAFVDFKARKQCLWFPFNSSSGVKKFGEHFD 86

Db 37 TKKQLGAGSIECAAKCEDE--EFTCRAFOYHSKEQOCVMAENRKSIIIRM--RDVV 92

Qy 87 LYENKDYIRNCLIGKRSYKGTVTSIKSGIKQCPWSSMIPHEHSFLPSSYRGKDLQENYC 146

Db 93 LFEKKVYLSECKTGNGKRYGTMTSKNGITCQKWSSTSPHRPRFSPATHPSEGLEENYC 152

Qy 147 RNPGEEGGPMCFSTSNPEVRYEVCIDIPOCSEVECMTCNGESYRGLMDHTESGKICQWDH 206

Db 153 RNPNDPOGPMCYITDPKRYDYCDILECEE--ECMHCSGENYDGKISKTMSGLECQAWDS 211

Qy 207 QTPHRHKFLPERYPDKGFDNNCRNPDQGPWCYTLDPHTRWEYCAIKTCADNTMNDTD 266

Db 212 QSPHAGYIPSKFKNLKNKYNCRNPDRELPRWCFTTDPNKRKWLCDIPRCT-----TP 265

Qy 267 VPLE--TTECTOGGEGYRGVNTIWIWICPQRWDSQYPHEHDMTPENFKCKDLRENVCYR 324

Db 266 PSSGPTYQCLUGTGNGYRGNAVTVSGHTCOHWSAQTPTHNRTPENFPCKNLDENYCR 325

Qy 325 NPDGSESPWCFTTDPNIRVGYCSQIPNCDMSH-----GDCYRGNGKN 367

Db 326 NPDGRAPWCHTTNSQVRWEYC-KIPSCDSSPVSTEQALPTAPPELTVPVQDCYHGDQS 384

Qy 368 YMGNLSTQSRGLTCSMDKNMEDLHRHIFWE---PDASKLNEYCRNPDHDAHGPWCYTG 424

Db 385 YRGTSSTTTTKKCKQSSWMT--HRHOKTPENYPNAG-LTMNYCRNPDAD-KGPWCFTT 440

Qy 425 NPLIPWDYCPISRCEDTPTIV 447

Db 441 DPSVRWEYCNLKKCSG-TEASVV 462

RESULT 13

US-09-946-893-8

; Sequence 8, Application US/09946893

; Patent No. US20020072494A1

; GENERAL INFORMATION:

; APPLICANT: Cao, Yihai

; TITLE OF INVENTION: Materials and methods relating to endothelial cell growth

; FILE REFERENCE: Mewburn

; CURRENT APPLICATION NUMBER: US/09/946,893

; PRIOR FILING DATE: 2001-09-05

; PRIOR FILING DATE: 2000-09-05

; NUMBER OF SEQ ID NOS: 11

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 8

; LENGTH: 571

; TYPE: PRT

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Description of Artificial Sequence: Angioquiescin  
US-09-946-893-8

Query Match

Best Local Similarity 43.6%; Pred. No. 4.3e-76;

Matches 193; Conservative 68; Mismatches 140; Indels 42; Gaps 13;

Qy 30 TKK---VNTADQCANRTRNKGLPFTCKAFVDFKARKQCLWFPFNSSSGVKKFGEHFD 86

Db 37 TKKQLGAGSIECAAKCEDE--EFTCRAFOYHSKEQOCVMAENRKSIIIRM--RDVV 92

Qy 87 LYENKDYIRNCLIGKRSYKGTVTSIKSGIKQCPWSSMIPHEHSFLPSSYRGKDLQENYC 146

Db 93 LFEKKVYLSECKTGNGKRYGTMTSKNGITCQKWSSTSPHRPRFSPATHPSEGLEENYC 152

Qy 147 RNPGEEGGPMCFSTSNPEVRYEVCIDIPOCSEVECMTCNGESYRGLMDHTESGKICQWDH 206

Db 153 RNPNDPOGPMCYITDPKRYDYCDILECEE--ECMHCSGENYDGKISKTMSGLECQAWDS 211

Qy 207 QTPHRHKFLPERYPDKGFDNNCRNPDQGPWCYTLDPHTRWEYCAIKTCADNTMNDTD 266

Db 212 QSPHAGYIPSKFKNLKNKYNCRNPDRELPRWCFTTDPNKRKWLCDIPRCT-----TP 265

Qy 267 VPLE--TTECTOGGEGYRGVNTIWIWICPQRWDSQYPHEHDMTPENFKCKDLRENVCYR 324

Db 266 PSSGPTYQCLUGTGNGYRGNAVTVSGHTCOHWSAQTPTHNRTPENFPCKNLDENYCR 325

Qy 325 NPDGSESPWCFTTDPNIRVGYCSQIPNCDMSH-----GDCYRGNGKN 367

Db 326 NPDGRAPWCHTTNSQVRWEYC-KIPSCDSSPVSTEQALPTAPPELTVPVQDCYHGDQS 384

Qy 368 YMGNLSTQSRGLTCSMDKNMEDLHRHIFWE---PDASKLNEYCRNPDHDAHGPWCYTG 424

Db 385 YRGTSSTTTTKKCKQSSWMT--HRHOKTPENYPNAG-LTMNYCRNPDAD-KGPWCFTT 440

Qy 425 NPLIPWDYCPISRCEDTPTIV 447

Db 441 DPSVRWEYCNLKKCSG-TEASVV 462

RESULT 14

US-09-946-893-6

; Sequence 6, Application US/09946893

; Patent No. US20020072494A1

; GENERAL INFORMATION:

; APPLICANT: Cao, Yihai

; TITLE OF INVENTION: Materials and methods relating to endothelial cell growth

; FILE REFERENCE: Mewburn

; CURRENT APPLICATION NUMBER: US/09/946,893

; PRIOR FILING DATE: 2001-09-05

; PRIOR FILING DATE: 2000-09-05

; NUMBER OF SEQ ID NOS: 11

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 6

; LENGTH: 576

; TYPE: PRT

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Description of Artificial Sequence: Angioquiescin  
US-09-946-893-6

Query Match

Best Local Similarity 43.6%; Pred. No. 4.3e-76;

Matches 193; Conservative 68; Mismatches 140; Indels 42; Gaps 13;

Qy 30 TKK---VNTADQCANRTRNKGLPFTCKAFVDFKARKQCLWFPFNSSSGVKKFGEHFD 86

Db 37 TKKQLGAGSIECAAKCEDE--EFTCRAFOYHSKEQOCVMAENRKSIIIRM--RDVV 92

Qy 87 LYENKDYIRNCLIGKRSYKGTVTSIKSGIKQCPWSSMIPHEHSFLPSSYRGKDLQENYC 146

Db 93 LFEKKVYLSECKTGNGKRYGTMTSKNGITCQKWSSTSPHRPRFSPATHPSEGLEENYC 152

Qy 147 RNPGEEGGPMCFSTSNPEVRYEVCIDIPOCSEVECMTCNGESYRGLMDHTESGKICQWDH 206

Db 153 RNPNDPOGPMCYITDPKRYDYCDILECEE--ECMHCSGENYDGKISKTMSGLECQAWDS 211

Qy	368	YMGNLQTSRGLTCSWMDKNMEDLHRHFWF----	PDASKLNYCYCRNPDDDAHGHCWYTG	420
Db	366	YRGTSTTTTGKKQCSWSSWTP--	HRHQKTPENYVAG-LIMNYCRNPDAD-KGFWCFET	421
Qy	425	NPLIPWDYCPISRCEGDTTFIV	447	
		:   :   :   :   :   :   :   :   :		
Db	422	DPSVRWEYCNLKKCSG-TEASVY	443	

Search completed: April 25, 2003, 15:49:03  
Job time : 25 secs

[illegible]

```

RESULT 15
US-09-967-386-1
; Sequence 1, Application US/09967386
; Patent No. US2002015992A1
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Henkin, Jack
; APPLICANT: Davidson, Donald J.
; TITLE OF INVENTION: ANTIAGIOGENIC POLYPEPTIDES AND METHODS
; TITLE OF INVENTION: FOR INHIBITING ANGIOGENESIS
; FILE REFERENCE: 6738 US.02
; CURRENT APPLICATION NUMBER: US/09/967,386
; CURRENT FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: US 60/236,550
; PRIOR FILING DATE: 2000-09-29
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 791
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-967-386-1

```

3

GenCore version 5.1.4\_p5.4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 25, 2003, 15:40:48 ; Search time 145 Seconds  
(without alignments)  
1987.557 Million cell updates/sec

Title: US-09-674-377B-1

Perfect score: 2604

Sequence: 1 ERKRRNTIHEFKKSAKTLI.....IPWDYCPISRCGGDTTPTIV 447

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 4569144 seqs, 644733110 residues

Total number of hits satisfying chosen parameters: 4569144

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending\_Patents\_AA\_Main.\*

```

1: /cgn2_6/ptodata/1/paa/PCRUS_COMB.pep.*
2: /cgn2_6/ptodata/1/paa/US06_COMB.pep.*
3: /cgn2_6/ptodata/1/paa/US07_COMB.pep.*
4: /cgn2_6/ptodata/1/paa/US080_COMB.pep.*
5: /cgn2_6/ptodata/1/paa/US081_COMB.pep.*
6: /cgn2_6/ptodata/1/paa/US082_COMB.pep.*
7: /cgn2_6/ptodata/1/paa/US083_COMB.pep.*
8: /cgn2_6/ptodata/1/paa/US084_COMB.pep.*
9: /cgn2_6/ptodata/1/paa/US085_COMB.pep.*
10: /cgn2_6/ptodata/1/paa/US086_COMB.pep.*
11: /cgn2_6/ptodata/1/paa/US087_COMB.pep.*
12: /cgn2_6/ptodata/1/paa/US088_COMB.pep.*
13: /cgn2_6/ptodata/1/paa/US089_COMB.pep.*
14: /cgn2_6/ptodata/1/paa/US090_COMB.pep.*
15: /cgn2_6/ptodata/1/paa/US091_COMB.pep.*
16: /cgn2_6/ptodata/1/paa/US092_COMB.pep.*
17: /cgn2_6/ptodata/1/paa/US093_COMB.pep.*
18: /cgn2_6/ptodata/1/paa/US094_COMB.pep.*
19: /cgn2_6/ptodata/1/paa/US095_COMB.pep.*
20: /cgn2_6/ptodata/1/paa/US096_COMB.pep.*
21: /cgn2_6/ptodata/1/paa/US097_COMB.pep.*
22: /cgn2_6/ptodata/1/paa/US098_COMB.pep.*
23: /cgn2_6/ptodata/1/paa/US099_COMB.pep.*
24: /cgn2_6/ptodata/1/paa/US100_COMB.pep.*
25: /cgn2_6/ptodata/1/paa/US101_COMB.pep.*
26: /cgn2_6/ptodata/1/paa/US102_COMB.pep.*
27: /cgn2_6/ptodata/1/paa/US60_COMB.pep.*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2604	100.0	447	20	US-09-674-377B-1
2	2601	99.9	447	23	US-09-951-629-1
3	2601	99.9	447	24	US-10-081-309-2
4	2601	99.9	697	5	US-08-167-641A-44
5	2601	99.9	697	8	US-08-458-217-50
6	2601	99.9	697	8	US-08-460-890-44

```

7 2601 99.9 697 8 US-08-460-971-50
8 2601 99.9 728 3 US-07-705-741-1
9 2601 99.9 728 3 US-07-712-932-3
10 2601 99.9 728 3 US-07-983-366-4
11 2601 99.9 728 6 US-08-242-802-2
12 2601 99.9 728 6 US-08-292-160B-1
13 2601 99.9 728 10 US-08-605-221-2
14 2601 99.9 728 13 US-08-909-515-2
15 2601 99.9 728 13 US-08-930-999-1
16 2601 99.9 728 18 US-09-423-516-2
17 2601 99.9 728 21 US-09-791-537-140648
18 2592 99.5 728 13 US-08-963-643-1
19 2566.5 98.6 442 20 US-09-674-377B-2
20 2565 98.5 728 21 US-09-791-537-36146
21 2565 98.5 728 21 US-09-791-537-140804
22 2565 98.5 728 27 US-60-382-445-2889
23 2563.5 98.4 487 20 US-09-601-040A-27
24 2563.5 98.4 717 20 US-09-601-040-6
25 2563.5 98.4 717 20 US-09-601-040A-6
26 2563.5 98.4 717 20 US-09-601-040A-8
27 2563.5 98.4 723 3 US-07-712-932-5
28 2563.5 98.4 723 6 US-08-267-837-4
29 2563.5 98.4 723 6 US-08-290-937-1
30 2563.5 98.4 723 6 US-08-290-937-2
31 2563.5 98.4 723 10 US-08-605-221-4
32 2563.5 98.4 723 20 US-09-600-991-18
33 2563.5 98.4 723 20 US-09-601-040-10
34 2563.5 98.4 723 20 US-09-601-040A-10
35 2563.5 98.4 723 21 US-09-791-537-140107
36 2563.5 98.4 729 20 US-09-601-040-2
37 2563.5 98.4 729 20 US-09-601-040-4
38 2563.5 98.4 729 20 US-09-601-040A-2
39 2563.5 98.4 729 20 US-09-601-040A-4
40 2563.5 98.4 723 3 US-07-705-741-2
41 2557.5 98.2 723 6 US-08-290-937-3
42 2557.5 98.2 723 6 US-08-292-160B-2
43 2557.5 98.2 836 27 US-60-230-435-1887
44 2545 97.6 723 25 US-10-133-912-18
45 2541.5

```

#### ALIGNMENTS

```

RESULT 1
US-09-674-377B-1
; Sequence 1, Application US/09674377B
; GENERAL INFORMATION:
; APPLICANT: Nakamura, Toshikazu
; TITLE OF INVENTION: NEOVASCULARIZATION INHIBITORS
; FILE REFERENCE: Q61434
; CURRENT APPLICATION NUMBER: US/09/674,377B
; CURRENT FILING DATE: 2000-10-30
; PRIOR APPLICATION NUMBER: PCT/JF99/01834
; PRIOR FILING DATE: 1999-04-06
; PRIOR APPLICATION NUMBER: JP 1998-134681
; PRIOR FILING DATE: 1998-04-28
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 447
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: MOD_RES
; LOCATION: (1)..(1)
; OTHER INFORMATION: pyroglutamate
; FEATURE:
; NAME/KEY: CHAIN
; LOCATION: (1)..(447)
; OTHER INFORMATION: N-terminal region of alpha-chain in HGF (PyrGlu32-val478/HGF)
US-09-674-377B-1

```

Sequence 50, Appli  
Sequence 1, Appli  
Sequence 3, Appli  
Sequence 4, Appli  
Sequence 2, Appli  
Sequence 2, Appli  
Sequence 2, Appli  
Sequence 140648,  
Sequence 1, Appli  
Sequence 2, Appli  
Sequence 36146, A  
Sequence 140804,  
Sequence 2889, Ap  
Sequence 27, Appli  
Sequence 6, Appli  
Sequence 8, Appli  
Sequence 6, Appli  
Sequence 8, Appli  
Sequence 5, Appli  
Sequence 4, Appli  
Sequence 1, Appli  
Sequence 2, Appli  
Sequence 4, Appli  
Sequence 2, Appli  
Sequence 2, Appli  
Sequence 140107,  
Sequence 4, Appli  
Sequence 2, Appli  
Sequence 4, Appli  
Sequence 2, Appli  
Sequence 2, Appli  
Sequence 3, Appli  
Sequence 1887, Ap  
Sequence 18, Appli

```

Query Match      100.0%; Score 2604; DB 20; Length 447;
Best Local Similarity 100.0%; Pred. No. 2.9e-230;
Matches 447; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ERKRRNTIHEFKSAKTTLIKIDPALKIKTKKVNADQCANRCTRNKGLPFTCKAFVFDK 60
DB 1 ERKRRNTIHEFKSAKTTLIKIDPALKIKTKKVNADQCANCTRNKGLPFTCKAFVFDK 60
QY 61 ARKQCLWFFPNSMSGVKKFEGHFDLYENKDYIRNCIIIGKRSYKGTVSIITKSGIKCQP 120
DB 61 ARKQCLWFFPNSMSGVKKFEGHFDLYENKDYIRNCIIIGKRSYKGTVSIITKSGIKCQP 120
QY 121 WSSMIPHEHSFLPSSYRGKDLQENYCRNPRGEGGPGWFTSNPEVRYEVCIDIPOCSEVEC 180
DB 121 WSSMIPHEHSFLPSSYRGKDLQENYCRNPRGEGGPGWFTSNPEVRYEVCIDIPOCSEVEC 180
QY 181 MTCNGESYRGLMDHTESGKICQWRDHTQPHRHKFLPERYDPKGFDDNYCRNPDGQPRPWC 240
DB 181 MTCNGESYRGLMDHTESGKICQWRDHTQPHRHKFLPERYDPKGFDDNYCRNPDGQPRPWC 240
QY 241 YTLDPHTRWEYCAIKTCADNTMNDTVPLETTECIOGQEGYRGTVNTIWNIPQCRWDS 300
DB 241 YTLDPHTRWEYCAIKTCADNTMNDTVPLETTECIOGQEGYRGTVNTIWNIPQCRWDS 300
QY 301 QYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDMSHGQDC 360
DB 301 QYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDMSHGQDC 360
QY 361 YRNGKNTMGNLSQTRSGLTCSMWDKNMEDLHRHIFWEPDASKNLNENYCRNPDGQPRPWC 420
DB 361 YRNGKNTMGNLSQTRSGLTCSMWDKNMEDLHRHIFWEPDASKNLNENYCRNPDGQPRPWC 420
QY 421 CYTGNPLIPWDYCPISRCGDTTPTIV 447
DB 421 CYTGNPLIPWDYCPISRCGDTTPTIV 447

```

```

RESULT 2
US-09-951-629-1
; Sequence 1, Application US/09951629
; GENERAL INFORMATION:
; APPLICANT: NAKAMURA, Toshikazu
; TITLE OF INVENTION: Anti-Cancer Agent
; FILE REFERENCE: 4373-0101P
; CURRENT APPLICATION NUMBER: US/09/951,629
; CURRENT FILING DATE: 2001-09-14
; PRIOR APPLICATION NUMBER: JP 300728
; PRIOR FILING DATE: 1995-10-24
; PRIOR APPLICATION NUMBER: PCT/JP96/03105
; PRIOR FILING DATE: 1996-10-23
; PRIOR APPLICATION NUMBER: US 09/471,032
; PRIOR FILING DATE: 1999-12-23
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 1
; LENGTH: 447
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-951-629-1

```

```

Query Match      99.9%; Score 2601; DB 23; Length 447;
Best Local Similarity 99.8%; Pred. No. 5.5e-230;
Matches 446; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ERKRRNTIHEFKSAKTTLIKIDPALKIKTKKVNADQCANRCTRNKGLPFTCKAFVFDK 60
DB 1 ERKRRNTIHEFKSAKTTLIKIDPALKIKTKKVNADQCANCTRNKGLPFTCKAFVFDK 60
QY 61 ARKQCLWFFPNSMSGVKKFEGHFDLYENKDYIRNCIIIGKRSYKGTVSIITKSGIKCQP 120
DB 61 ARKQCLWFFPNSMSGVKKFEGHFDLYENKDYIRNCIIIGKRSYKGTVSIITKSGIKCQP 120
QY 121 WSSMIPHEHSFLPSSYRGKDLQENYCRNPRGEGGPGWFTSNPEVRYEVCIDIPOCSEVEC 180

```

```

DB 121 WSSMIPHEHSFLPSSYRGKDLQENYCRNPRGEGGPGWFTSNPEVRYEVCIDIPOCSEVEC 180
QY 181 MTCNGESYRGLMDHTESGKICQWRDHTQPHRHKFLPERYDPKGFDDNYCRNPDGQPRPWC 240
DB 181 MTCNGESYRGLMDHTESGKICQWRDHTQPHRHKFLPERYDPKGFDDNYCRNPDGQPRPWC 240
QY 241 YTLDPHTRWEYCAIKTCADNTMNDTVPLETTECIOGQEGYRGTVNTIWNIPQCRWDS 300
DB 241 YTLDPHTRWEYCAIKTCADNTMNDTVPLETTECIOGQEGYRGTVNTIWNIPQCRWDS 300
QY 301 QYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDMSHGQDC 360
DB 301 QYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDMSHGQDC 360
QY 361 YRNGKNTMGNLSQTRSGLTCSMWDKNMEDLHRHIFWEPDASKNLNENYCRNPDGQPRPWC 420
DB 361 YRNGKNTMGNLSQTRSGLTCSMWDKNMEDLHRHIFWEPDASKNLNENYCRNPDGQPRPWC 420
QY 421 CYTGNPLIPWDYCPISRCGDTTPTIV 447
DB 421 CYTGNPLIPWDYCPISRCGDTTPTIV 447

```

```

RESULT 3
US-10-081-309-2
; Sequence 2, Application US/10081309
; GENERAL INFORMATION:
; APPLICANT: Hoffmann-La Roche Inc.
; TITLE OF INVENTION: PEG Conjugates of NK4
; FILE REFERENCE: 20859
; CURRENT APPLICATION NUMBER: US/10/081,309
; CURRENT FILING DATE: 2002-06-04
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 2
; LENGTH: 447
; TYPE: PRT
; ORGANISM: homosapiens
US-10-081-309-2

```

```

Query Match      99.9%; Score 2601; DB 24; Length 447;
Best Local Similarity 99.8%; Pred. No. 5.5e-230;
Matches 446; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ERKRRNTIHEFKSAKTTLIKIDPALKIKTKKVNADQCANCTRNKGLPFTCKAFVFDK 60
DB 1 ERKRRNTIHEFKSAKTTLIKIDPALKIKTKKVNADQCANCTRNKGLPFTCKAFVFDK 60
QY 61 ARKQCLWFFPNSMSGVKKFEGHFDLYENKDYIRNCIIIGKRSYKGTVSIITKSGIKCQP 120
DB 61 ARKQCLWFFPNSMSGVKKFEGHFDLYENKDYIRNCIIIGKRSYKGTVSIITKSGIKCQP 120
QY 121 WSSMIPHEHSFLPSSYRGKDLQENYCRNPRGEGGPGWFTSNPEVRYEVCIDIPOCSEVEC 180
DB 121 WSSMIPHEHSFLPSSYRGKDLQENYCRNPRGEGGPGWFTSNPEVRYEVCIDIPOCSEVEC 180
QY 181 MTCNGESYRGLMDHTESGKICQWRDHTQPHRHKFLPERYDPKGFDDNYCRNPDGQPRPWC 240
DB 181 MTCNGESYRGLMDHTESGKICQWRDHTQPHRHKFLPERYDPKGFDDNYCRNPDGQPRPWC 240
QY 241 YTLDPHTRWEYCAIKTCADNTMNDTVPLETTECIOGQEGYRGTVNTIWNIPQCRWDS 300
DB 241 YTLDPHTRWEYCAIKTCADNTMNDTVPLETTECIOGQEGYRGTVNTIWNIPQCRWDS 300
QY 301 QYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDMSHGQDC 360
DB 301 QYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDMSHGQDC 360
QY 361 YRNGKNTMGNLSQTRSGLTCSMWDKNMEDLHRHIFWEPDASKNLNENYCRNPDGQPRPWC 420
DB 361 YRNGKNTMGNLSQTRSGLTCSMWDKNMEDLHRHIFWEPDASKNLNENYCRNPDGQPRPWC 420

```

QY 421 CYTGNPLIPWDYCPISRCGDTTPTIV 447  
 Db 421 CYTGNPLIPWDYCPISRCGDTTPTIV 447

## RESULT 4

US-08-167-641A-44  
 ; Sequence 44, Application US/08167641A  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Woo, Savio L.C.  
 ; APPLICANT: Smith, Louis C.  
 ; APPLICANT: Cristiano, Richard J.  
 ; APPLICANT: Gottchalk, Stephen  
 ; TITLE OF INVENTION: NUCLEIC ACID TRANSPORTER  
 ; TITLE OF INVENTION: SYSTEMS AND METHODS OF  
 ; TITLE OF INVENTION: USE  
 ; NUMBER OF SEQUENCES: 51  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: LYON & LYON  
 ; STREET: 611 West Sixth St.  
 ; CITY: Los Angeles  
 ; STATE: California  
 ; COUNTRY: U.S.A.  
 ; ZIP: 90017  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
 ; MEDIUM TYPE: storage  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: IBM MS-DOS (Ver. 5.0)  
 ; SOFTWARE: WordPerfect (Ver. 5.1)  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/167,641A  
 ; FILING DATE: December 14, 1993  
 ; CLASSIFICATION: 435  
 ; PRIOR APPLICATION DATA:  
 ; PRIOR APPLICATION DATA: including application  
 ; PRIOR APPLICATION DATA: described below:  
 ; APPLICATION NUMBER: 07/855,389  
 ; FILING DATE: March 20, 1992  
 ; APPLICATION NUMBER: PCT/US93/02725  
 ; FILING DATE: March 19, 1993  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Knight, Matthew W.  
 ; REGISTRATION NUMBER: 36,846  
 ; REFERENCE/DOCKET NUMBER: 205/012  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (213) 489-1600  
 ; TELEFAX: (213) 955-0440  
 ; TELEX: 67-3510  
 ; INFORMATION FOR SEQ ID NO: 44:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 697 amino acids  
 ; TYPE: amino acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: protein  
 ; US-08-167-641A-44

Query Match 99.9%; Score 2601; DB 5; Length 697;  
 Best Local Similarity 99.8%; Pred No. 1e-229;  
 Matches 446; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ERKRRTTHFEKSAKTLIKIDPALKIKTKVNTADQCANRCTRNGKGLPTCKAFVDEK 60  
 Db 1 QKRKRRTTHFEKSAKTLIKIDPALKIKTKVNTADQCANRCTRNGKGLPTCKAFVDEK 60

QY 61 ARKQCLWFPNMSGKVEGHEFDLYENKDYTRNCIIIGKRSYKGTVSTKSGIKCQP 120  
 Db 61 ARKQCLWFPNMSGKVEGHEFDLYENKDYTRNCIIIGKRSYKGTVSTKSGIKCQP 120

QY 121 WSSMIPHEHFLPSSYRGKDLQENYCRNPRGEGGWCFTSNPEVRYEVCIDIPOCSEVEC 180  
 Db 121 WSSMIPHEHFLPSSYRGKDLQENYCRNPRGEGGWCFTSNPEVRYEVCIDIPOCSEVEC 180

QY 181 MTCNGESYRGLMDHTESGKICQRWDHOTPHRKELPERYPDKGFDNDYCRNPDGQPREWC 240  
 Db 181 MTCNGESYRGLMDHTESGKICQRWDHOTPHRKELPERYPDKGFDNDYCRNPDGQPREWC 240

QY 241 YTLDPHTRWEYCAIKTCADNTMDTDPVLETTECICQGGEGYRGTVNTIWNIGIPCQRWDS 300  
 Db 241 YTLDPHTRWEYCAIKTCADNTMDTDPVLETTECICQGGEGYRGTVNTIWNIGIPCQRWDS 300

QY 301 QYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDMSHGDC 360  
 Db 301 QYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDMSHGDC 360

QY 361 YRGNGKNYMGNSQTRSLGTCMMDKNMDLHRHIFWEPDASKLKNENYCRNPDGDDAHGPW 420  
 Db 361 YRGNGKNYMGNSQTRSLGTCMMDKNMDLHRHIFWEPDASKLKNENYCRNPDGDDAHGPW 420

QY 421 CYTGNPLIPWDYCPISRCGDTTPTIV 447  
 Db 421 CYTGNPLIPWDYCPISRCGDTTPTIV 447

## RESULT 5

US-08-458-217-50  
 ; Sequence 50, Application US/08458217  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Woo, Savio L.C.  
 ; APPLICANT: Smith, Louis C.  
 ; APPLICANT: Cristiano, Richard J.  
 ; APPLICANT: Gottchalk, Stephen  
 ; TITLE OF INVENTION: NUCLEIC ACID TRANSPORTER SYSTEMS AND  
 ; TITLE OF INVENTION: METHODS OF USE  
 ; NUMBER OF SEQUENCES: 65  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Lyon & Lyon  
 ; STREET: 633 West Fifth Street  
 ; CITY: Los Angeles  
 ; STATE: California  
 ; COUNTRY: U.S.A.  
 ; ZIP: 90071-2066  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
 ; MEDIUM TYPE: storage  
 ; COMPUTER: IBM Compatible  
 ; OPERATING SYSTEM: IBM P.C. DOS 5.0  
 ; SOFTWARE: FastSeq for Windows 2.0  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/458,217  
 ; FILING DATE: June 5, 1995  
 ; CLASSIFICATION: 514  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: 08/167,641  
 ; FILING DATE: December 14, 1993  
 ; APPLICATION NUMBER: 07/855,389  
 ; FILING DATE: March 20, 1992  
 ; APPLICATION NUMBER: PCT/US93/02725  
 ; FILING DATE: March 19, 1993  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Warburg, Richard J.  
 ; REGISTRATION NUMBER: 32,327  
 ; REFERENCE/DOCKET NUMBER: 212/065  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (213) 489-1600  
 ; TELEFAX: (213) 955-0440  
 ; TELEX: 67-3510  
 ; INFORMATION FOR SEQ ID NO: 50:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 697 amino acids  
 ; TYPE: amino acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: peptide

Sat Apr 26 11:05:57 2003

us-09-674-377b-1.rapm

US-08-458-217-50

Query Match 99.9%; Score 2601; DB 8; Length 697;  
Best Local Similarity 99.8%; Pred. No. 1e-229;  
Matches 446; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 ERKRRNTIHEFKSAKTLIKIDPALKIKTKVNTADOCANRTRNKGLPFTCKAFVFDK 60
DB 1 QKRRNTIHEFKSAKTLIKIDPALKIKTKVNTADOCANRTRNKGLPFTCKAFVFDK 60
QY 61 ARKQCLWFPNMSGVKKKEGHEFDLYENKDYIRNCIIIGKRSYKGTVSITKSGIKCQP 120
DB 61 ARKQCLWFPNMSGVKKKEGHEFDLYENKDYIRNCIIIGKRSYKGTVSITKSGIKCQP 120
QY 121 WSSMIPHEHSLPSSYRGKDLQENYCRNPRGEGGWCFTSNPEVRYEVCIDIQCSEVEC 180
DB 121 WSSMIPHEHSLPSSYRGKDLQENYCRNPRGEGGWCFTSNPEVRYEVCIDIQCSEVEC 180
QY 181 MTCNGESYRGLMDHTESGKICQWDHOTPHRHKFLPERYPDKGDDNYCRNPDGQPRPWC 240
DB 181 MTCNGESYRGLMDHTESGKICQWDHOTPHRHKFLPERYPDKGDDNYCRNPDGQPRPWC 240
QY 241 YLDPHTRWYCAIKTCADNTMNDTVDPLETTCIOGEGYRGTVNTIWNIGPCQRWDS 300
DB 241 YLDPHTRWYCAIKTCADNTMNDTVDPLETTCIOGEGYRGTVNTIWNIGPCQRWDS 300
QY 301 QYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDMSHGQDC 360
DB 301 QYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDMSHGQDC 360
QY 361 YRGNGKNYMGNLSTQTRSGLTCSMDKNMEDLHRHIFWEPDASKLKNENYCRNPDDDAHGPW 420
DB 361 YRGNGKNYMGNLSTQTRSGLTCSMDKNMEDLHRHIFWEPDASKLKNENYCRNPDDDAHGPW 420
QY 421 CYTGNPLIPWDYCPISRCGEGDTTPTIV 447
DB 421 CYTGNPLIPWDYCPISRCGEGDTTPTIV 447
```

RESULT 6

US-08-460-890-44

Sequence 44, Application US/08460890  
GENERAL INFORMATION:  
APPLICANT: Woo, Savio L.C.  
APPLICANT: Smith, Louis C.  
APPLICANT: Cristiano, Richard J.  
APPLICANT: Gottchalk, Stephen  
TITLE OF INVENTION: NUCLEIC ACID TRANSPORTER  
TITLE OF INVENTION: SYSTEMS AND METHODS OF  
TITLE OF INVENTION: USE  
NUMBER OF SEQUENCES: 51  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: LYON & LYON  
STREET: 633 West Fifth St.  
CITY: Los Angeles  
STATE: California  
COUNTRY: U.S.A.  
ZIP: 90071

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
MEDIUM TYPE: storage  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: IBM MS-DOS (Ver. 6.22)  
SOFTWARE: Wordperfect (Ver. 5.1)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/460,890  
FILING DATE: June 5, 1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA: including application  
PRIOR APPLICATION DATA: described below:  
APPLICATION NUMBER: 08/167,641  
FILING DATE: December 14, 1993

RESULT 7

US-08-460-971-50  
Sequence 50, Application US/08460971  
GENERAL INFORMATION:  
APPLICANT: Woo, Savio L.C.  
APPLICANT: Smith, Louis C.  
APPLICANT: Cristiano, Richard J.  
APPLICANT: Gottchalk, Stephen  
TITLE OF INVENTION: NUCLEIC ACID TRANSPORTER SYSTEMS AND  
TITLE OF INVENTION: METHODS OF USE  
NUMBER OF SEQUENCES: 65  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Lyon & Lyon  
STREET: 633 West Fifth Street  
STREET: Suite 4700

APPLICATION NUMBER: 07/855,389  
FILING DATE: March 20, 1992  
APPLICATION NUMBER: PCT/US93/02725  
FILING DATE: March 19, 1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Knight, Matthew W.  
REGISTRATION NUMBER: 36,846  
REFERENCE/DOCKET NUMBER: 212/066  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (213) 489-1600  
TELEFAX: (213) 955-0440  
TELEX: 67-3510  
INFORMATION FOR SEQ ID NO: 44:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 697 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-460-890-44

Query Match 99.9%; Score 2601; DB 8; Length 697;

Best Local Similarity 99.8%; Pred. No. 1e-229;  
Matches 446; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 ERKRRNTIHEFKSAKTLIKIDPALKIKTKVNTADOCANRTRNKGLPFTCKAFVFDK 60
DB 1 QKRRNTIHEFKSAKTLIKIDPALKIKTKVNTADOCANRTRNKGLPFTCKAFVFDK 60
QY 61 ARKQCLWFPNMSGVKKKEGHEFDLYENKDYIRNCIIIGKRSYKGTVSITKSGIKCQP 120
DB 61 ARKQCLWFPNMSGVKKKEGHEFDLYENKDYIRNCIIIGKRSYKGTVSITKSGIKCQP 120
QY 121 WSSMIPHEHSLPSSYRGKDLQENYCRNPRGEGGWCFTSNPEVRYEVCIDIQCSEVEC 180
DB 121 WSSMIPHEHSLPSSYRGKDLQENYCRNPRGEGGWCFTSNPEVRYEVCIDIQCSEVEC 180
QY 181 MTCNGESYRGLMDHTESGKICQWDHOTPHRHKFLPERYPDKGDDNYCRNPDGQPRPWC 240
DB 181 MTCNGESYRGLMDHTESGKICQWDHOTPHRHKFLPERYPDKGDDNYCRNPDGQPRPWC 240
QY 241 YLDPHTRWYCAIKTCADNTMNDTVDPLETTCIOGEGYRGTVNTIWNIGPCQRWDS 300
DB 241 YLDPHTRWYCAIKTCADNTMNDTVDPLETTCIOGEGYRGTVNTIWNIGPCQRWDS 300
QY 301 QYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDMSHGQDC 360
DB 301 QYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDMSHGQDC 360
QY 361 YRGNGKNYMGNLSTQTRSGLTCSMDKNMEDLHRHIFWEPDASKLKNENYCRNPDDDAHGPW 420
DB 361 YRGNGKNYMGNLSTQTRSGLTCSMDKNMEDLHRHIFWEPDASKLKNENYCRNPDDDAHGPW 420
QY 421 CYTGNPLIPWDYCPISRCGEGDTTPTIV 447
DB 421 CYTGNPLIPWDYCPISRCGEGDTTPTIV 447
```



CITY: Los Angeles  
STATE: California  
COUNTRY: U.S.A.  
ZIP: 90071-2066  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
MEDIUM TYPE: storage  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: IBM P.C. DOS 5.0  
SOFTWARE: FASTSEQ for Windows 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/460,971  
FILING DATE: June 5, 1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/167,641  
FILING DATE: December 14, 1993  
APPLICATION NUMBER: 07/855,389  
FILING DATE: March 20, 1992  
APPLICATION NUMBER: PCT/US93/02725  
FILING DATE: March 19, 1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Warburg, Richard J.  
REGISTRATION NUMBER: 32,327  
REFERENCE/DOCKET NUMBER: 212/063  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (213) 489-1600  
TELEFAX: (213) 955-0440  
TELEX: 67-3510  
INFORMATION FOR SEQ ID NO: 50:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 697 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-460-971-50

Query Match 99.9%; Score 2601; DB 8; Length 697;  
Best Local Similarity 99.8%; Pred. No. 1e-229;  
Matches 446; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ERKRNTHEFKSAKTLIKIDPALKIKTKKVNADQCANRCTRNKGLPTCKAFVFDK 60  
Db :|||||  
Qy 1 QKRRNTHEFKSAKTLIKIDPALKIKTKKVNADQCANRCTRNKGLPTCKAFVFDK 60  
Db :|||||  
Qy 61 ARKQCLWFPFNSMSGVKGKFEHFDLYENKDYIRNCIIIGKRSYKGTVSTKSGIKQCP 120  
Db :|||||  
Qy 61 ARKQCLWFPFNSMSGVKGKFEHFDLYENKDYIRNCIIIGKRSYKGTVSTKSGIKQCP 120  
Db :|||||  
Qy 121 WSSMIPHEHSLPSSYRGKDLQENYCRNPRGEGPWCFTSNPEVRYEVCIDIPOCSEVEC 180  
Db :|||||  
Qy 181 MTCNGESYRGLMDHTESGKICQWRDHTPHRHKFLPERYPDKGDDNYCRNPDGQPRWC 240  
Db :|||||  
Qy 181 MTCNGESYRGLMDHTESGKICQWRDHTPHRHKFLPERYPDKGDDNYCRNPDGQPRWC 240  
Db :|||||  
Qy 241 YLDPHTRWECYCAIKTCADNTMNDTVDPLETTECIGOGEGYRGVTNTWNGIPCORWDS 300  
Db :|||||  
Qy 241 YLDPHTRWECYCAIKTCADNTMNDTVDPLETTECIGOGEGYRGVTNTWNGIPCORWDS 300  
Db :|||||  
Qy 301 QYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDMSHGQDC 360  
Db :|||||  
Qy 301 QYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDMSHGQDC 360  
Db :|||||  
Qy 361 YRGCKNTYNGNLSTQSLTCSMDKKNMEDLHRHIFWEPDASKLNEYCRNPDGDDAHGPW 420  
Db :|||||  
Qy 361 YRGCKNTYNGNLSTQSLTCSMDKKNMEDLHRHIFWEPDASKLNEYCRNPDGDDAHGPW 420  
Db :|||||  
Qy 421 CYTGNPLIPWDYCPISRCEGDTTPTIV 447  
Db :|||||  
Qy 421 CYTGNPLIPWDYCPISRCEGDTTPTIV 447  
Db :|||||

## RESULT 8

US-07-705-741-1  
Sequence 1, Application US/07705741  
GENERAL INFORMATION:  
APPLICANT: Nakamura, Toshikazu  
APPLICANT: Seki, Tatsuya  
APPLICANT: Hagiya, Michio  
APPLICANT: Shimomishi, Manabu  
APPLICANT: Shimizu, Shin  
TITLE OF INVENTION: Chromosome DNA Coding For Human  
TITLE OF INVENTION: Hepatocyte Growth Factor  
NUMBER OF SEQUENCES: 17  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Leydig, Voit & Mayer  
STREET: Two Prudential Plaza, Suite 4900  
CITY: Chicago  
STATE: IL  
COUNTRY: U.S.A.  
ZIP: 60601-6780  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07705,741  
FILING DATE: 19910528  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Green, Robert F.  
REGISTRATION NUMBER: 27,555  
REFERENCE/DOCKET NUMBER: 38300  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 312-616-5600  
TELEFAX: 312-616-5700  
TELEX: 25-3533  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 728 amino acids  
TYPE: AMINO ACID  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-07-705-741-1

Query Match 99.9%; Score 2601; DB 3; Length 728;  
Best Local Similarity 99.8%; Pred. No. 1e-229;  
Matches 446; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ERKRNTHEFKSAKTLIKIDPALKIKTKKVNADQCANRCTRNKGLPTCKAFVFDK 60  
Db :|||||  
Qy 32 QKRRNTHEFKSAKTLIKIDPALKIKTKKVNADQCANRCTRNKGLPTCKAFVFDK 91  
Db :|||||  
Qy 61 ARKQCLWFPFNSMSGVKGKFEHFDLYENKDYIRNCIIIGKRSYKGTVSTKSGIKQCP 120  
Db :|||||  
Qy 92 ARKQCLWFPFNSMSGVKGKFEHFDLYENKDYIRNCIIIGKRSYKGTVSTKSGIKQCP 151  
Db :|||||  
Qy 121 WSSMIPHEHSLPSSYRGKDLQENYCRNPRGEGPWCFTSNPEVRYEVCIDIPOCSEVEC 180  
Db :|||||  
Qy 152 WSSMIPHEHSLPSSYRGKDLQENYCRNPRGEGPWCFTSNPEVRYEVCIDIPOCSEVEC 211  
Db :|||||  
Qy 181 MTCNGESYRGLMDHTESGKICQWRDHTPHRHKFLPERYPDKGDDNYCRNPDGQPRWC 240  
Db :|||||  
Qy 212 MTCNGESYRGLMDHTESGKICQWRDHTPHRHKFLPERYPDKGDDNYCRNPDGQPRWC 271  
Db :|||||  
Qy 241 YLDPHTRWECYCAIKTCADNTMNDTVDPLETTECIGOGEGYRGVTNTWNGIPCORWDS 300  
Db :|||||  
Qy 272 YLDPHTRWECYCAIKTCADNTMNDTVDPLETTECIGOGEGYRGVTNTWNGIPCORWDS 331  
Db :|||||  
Qy 301 QYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDMSHGQDC 360  
Db :|||||  
Qy 332 QYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDMSHGQDC 391  
Db :|||||



Qy 1 ERKRNTTHERFKSAKTLIKIDPALKIKTKKVTADOCANCRTRNKGLPFTCKAFVDEK 60  
Db :|||||  
Qy 32 QKRKRNTTHERFKSAKTLIKIDPALKIKTKKVTADOCANCRTRNKGLPFTCKAFVDEK 91  
Db :|||||  
Qy 61 ARKQCLWFPFNSMSGVKKKEGHEFDLYENKDYIRNCIIIGKRSYKGTVSITKSGIKQCP 120  
Db :|||||  
Qy 92 ARKQCLWFPFNSMSGVKKKEGHEFDLYENKDYIRNCIIIGKRSYKGTVSITKSGIKQCP 151  
Db :|||||  
Qy 121 WSSMIPHEHSELPSSYRGKDLQENYCRNPRGEGGPGWCFSTNPEVRYEVCIDIPOCSEVEC 180  
Db :|||||  
Qy 152 WSSMIPHEHSELPSSYRGKDLQENYCRNPRGEGGPGWCFSTNPEVRYEVCIDIPOCSEVEC 211  
Db :|||||  
Qy 181 MTCNGESYRGLMDHTESGKICQWRDHQTPHRRKFLPERYPDKGDDNCRNPDGQPRWC 240  
Db :|||||  
Qy 212 MTCNGESYRGLMDHTESGKICQWRDHQTPHRRKFLPERYPDKGDDNCRNPDGQPRWC 271  
Db :|||||  
Qy 241 YLDPHTRWYCAIKTCADNTMNDTVPLETTTCIQGGEGYRGTVNTIWNIGIPCORWDS 300  
Db :|||||  
Qy 272 YLDPHTRWYCAIKTCADNTMNDTVPLETTTCIQGGEGYRGTVNTIWNIGIPCORWDS 331  
Db :|||||  
Qy 301 QYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDMSHGQDC 360  
Db :|||||  
Qy 332 QYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDMSHGQDC 391  
Db :|||||  
Qy 361 YRGNGKNTMGNSLQTRSGLTCSMDKNMEDLHRHIFWEPDASKLNENYCRNPDGQPRWC 420  
Db :|||||  
Qy 392 YRGNGKNTMGNSLQTRSGLTCSMDKNMEDLHRHIFWEPDASKLNENYCRNPDGQPRWC 451  
Db :|||||  
Qy 421 CYTGNPLIPWDYCPISRCGEGDTTPTIV 447  
Db :|||||  
Qy 452 CYTGNPLIPWDYCPISRCGEGDTTPTIV 478  
Db :|||||

## RESULT 11

US-08-242-802-2  
; Sequence 2, Application US/08242802  
; GENERAL INFORMATION:  
; APPLICANT: Nakamura, Toshikazu  
; TITLE OF INVENTION: Therapeutic Agent for Renal Disease  
; NUMBER OF SEQUENCES: 2  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Leydig, Voit & Mayer  
; STREET: Two Prudential Plaza, Suite 4900  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: U.S.A.  
; ZIP: 61801-6780  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/242,802  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/07/967,562  
; FILING DATE:  
; APPLICATION NUMBER: US 07/716,792  
; FILING DATE: 18-JUN-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Kilyk Jr., John  
; REGISTRATION NUMBER: 30,763  
; REFERENCE/DOCKET NUMBER: 38351  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 312-616-5600  
; TELEFAX: 312-616-5700  
; TELEX: 25-3533  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 728 amino acids  
; TYPE: amino acid

; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-242-802-2

Query Match 99.9%; Score 2601; DB 6; Length 728;  
Best Local Similarity 99.8%; Pred. No. 1.le-229;  
Matches 446; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 ERKRNTTHERFKSAKTLIKIDPALKIKTKKVTADOCANCRTRNKGLPFTCKAFVDEK 60  
Db :|||||  
Qy 32 QKRKRNTTHERFKSAKTLIKIDPALKIKTKKVTADOCANCRTRNKGLPFTCKAFVDEK 91  
Db :|||||  
Qy 61 ARKQCLWFPFNSMSGVKKKEGHEFDLYENKDYIRNCIIIGKRSYKGTVSITKSGIKQCP 120  
Db :|||||  
Qy 92 ARKQCLWFPFNSMSGVKKKEGHEFDLYENKDYIRNCIIIGKRSYKGTVSITKSGIKQCP 151  
Db :|||||  
Qy 121 WSSMIPHEHSELPSSYRGKDLQENYCRNPRGEGGPGWCFSTNPEVRYEVCIDIPOCSEVEC 180  
Db :|||||  
Qy 152 WSSMIPHEHSELPSSYRGKDLQENYCRNPRGEGGPGWCFSTNPEVRYEVCIDIPOCSEVEC 211  
Db :|||||  
Qy 181 MTCNGESYRGLMDHTESGKICQWRDHQTPHRRKFLPERYPDKGDDNCRNPDGQPRWC 240  
Db :|||||  
Qy 212 MTCNGESYRGLMDHTESGKICQWRDHQTPHRRKFLPERYPDKGDDNCRNPDGQPRWC 271  
Db :|||||  
Qy 241 YLDPHTRWYCAIKTCADNTMNDTVPLETTTCIQGGEGYRGTVNTIWNIGIPCORWDS 300  
Db :|||||  
Qy 272 YLDPHTRWYCAIKTCADNTMNDTVPLETTTCIQGGEGYRGTVNTIWNIGIPCORWDS 331  
Db :|||||  
Qy 301 QYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDMSHGQDC 360  
Db :|||||  
Qy 332 QYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDMSHGQDC 391  
Db :|||||  
Qy 361 YRGNGKNTMGNSLQTRSGLTCSMDKNMEDLHRHIFWEPDASKLNENYCRNPDGQPRWC 420  
Db :|||||  
Qy 392 YRGNGKNTMGNSLQTRSGLTCSMDKNMEDLHRHIFWEPDASKLNENYCRNPDGQPRWC 451  
Db :|||||  
Qy 421 CYTGNPLIPWDYCPISRCGEGDTTPTIV 447  
Db :|||||  
Qy 452 CYTGNPLIPWDYCPISRCGEGDTTPTIV 478  
Db :|||||

## RESULT 12

US-08-292-160B-1  
; Sequence 1, Application US/08292160B  
; GENERAL INFORMATION:  
; APPLICANT: Nakamura, Toshikazu  
; APPLICANT: Seki, Tatsuya  
; APPLICANT: Hagiya, Michio  
; APPLICANT: Shimonishi, Manabu  
; APPLICANT: Shimizu, Shin  
; TITLE OF INVENTION: Chromosome DNA Coding For Human  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; NUMBER OF SEQUENCES: 17  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Leydig, Voit & Mayer  
; STREET: Two Prudential Plaza, Suite 4900  
; CITY: Chicago  
; STATE: IL  
; COUNTRY: U.S.A.  
; ZIP: 60601-6780  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/292,160B  
; FILING DATE: 17-AUG-1994  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Green, Robert F.  
; REGISTRATION NUMBER: 27,555  
; REFERENCE/DOCKET NUMBER: 62748

```
TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-616-5600
; TELEFAX: 312-616-5700
; TELEX: 25-3533
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 728 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-292-160B-1

Query Match          99.9%; Score 2601; DB 6; Length 728;
Best Local Similarity 99.8%; Pred. No. 1.le-229;
Matches 446; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ERKRNTIHEFKSAKTTLIKIDPALKIKTKVNTADQCANCRNKGLPFTCKAFVFDK 60
Db :|||||
32 QKRRNTIHEFKSAKTTLIKIDPALKIKTKVNTADQCANCRNKGLPFTCKAFVFDK 91
Db :|||||
QY 61 ARKOCLWFPNSMSSGVKKEFGHEFDLYENKDYIRNCIIIGKGRSYKGTYSITKSGIKCOP 120
Db :|||||
QY 92 ARKOCLWFPNSMSSGVKKEFGHEFDLYENKDYIRNCIIIGKGRSYKGTYSITKSGIKCOP 151
Db :|||||
QY 121 WSSMIPHEHSFLPSSYRGKDLQENYCRNPRGEGGPMWCFSTSNPEVRYEYCDIPQCEVEEC 180
Db :|||||
QY 152 WSSMIPHEHSFLPSSYRGKDLQENYCRNPRGEGGPMWCFSTSNPEVRYEYCDIPQCEVEEC 211
Db :|||||
QY 181 MTCNGESYRGLMDHTESGKICORWDHOTPHRHKFLPERYDPKGFDDNYCRNPDGQPRPWC 240
Db :|||||
QY 212 MTCNGESYRGLMDHTESGKICORWDHOTPHRHKFLPERYDPKGFDDNYCRNPDGQPRPWC 271
Db :|||||
QY 241 YTLDPHTRWEYCAIKTCADNTMDTDVPLETTECIOGEGYRGTVNTIWNIGIPCORWDS 300
Db :|||||
QY 272 YTLDPHTRWEYCAIKTCADNTMDTDVPLETTECIOGEGYRGTVNTIWNIGIPCORWDS 331
Db :|||||
QY 301 QYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDMSHGQDC 360
Db :|||||
QY 332 QYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDMSHGQDC 391
Db :|||||
QY 361 YRGNGKNYMGNSLQTSRGLTCSMDKNMEDLHRHIFWEPDASKNLNENYCRNPDHDAHPW 420
Db :|||||
QY 392 YRGNGKNYMGNSLQTSRGLTCSMDKNMEDLHRHIFWEPDASKNLNENYCRNPDHDAHPW 451
Db :|||||
QY 421 CYTGNPLIPWDYCPISRCGDTTPTIV 447
Db :|||||
QY 452 CYTGNPLIPWDYCPISRCGDTTPTIV 478
Db :|||||

RESULT 14
US-08-909-515-2
; Sequence 2, Application US/08909515
; GENERAL INFORMATION:
; APPLICANT: KITAMURA, NAOMI
; APPLICANT: MIYAZAWA, KEIJI
; APPLICANT: DAIKOHARA, YASUSHI
; APPLICANT: TSUBOUCHI, HIROHITO
; APPLICANT: NAKA, DAIJA
; APPLICANT: TAKAHASHI, KAZUHIRO
; APPLICANT: MATSUI, RIE
; APPLICANT: YOSHIIYAMA, YOSHIKO
; APPLICANT: ISHII, TAKEHISA
; TITLE OF INVENTION: HEPATIC PARENCHYMAL CELL GROWTH FACTOR
; TITLE OF INVENTION: GENE ENCODING THE SAME, PROCESS
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN
; STREET: 130 WAKE STREET
; CITY: BOSTON
; STATE: MA
; COUNTRY: USA
; ZIP: 02019
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/909,515
; FILING DATE: 12-AUG-1997
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/455,986
; FILING DATE: 31-MAY-1995
; APPLICATION NUMBER: 08/089,417
; FILING DATE: 09-JUL-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/852,758
; FILING DATE: 17-MAR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/564,172

US-08-605-221-2
; Sequence 2, Application US/08605221
; GENERAL INFORMATION:
; APPLICANT: NAKAMURA, Toshikazu
; TITLE OF INVENTION: THERAPEUTIC AGENT FOR DISORDER IN BRAIN AND NERVE
; FILE REFERENCE: 2520-0101P
; CURRENT APPLICATION NUMBER: US/08/605,221
; CURRENT FILING DATE: 1996-03-15
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 728
; TYPE: PRT
; ORGANISM: Homo sapiens
; OTHER INFORMATION: HGF-A (Full length HGF)
US-08-605-221-2

Query Match          99.9%; Score 2601; DB 10; Length 728;
Best Local Similarity 99.8%; Pred. No. 1.le-229;
Matches 446; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ERKRNTIHEFKSAKTTLIKIDPALKIKTKVNTADQCANCRNKGLPFTCKAFVFDK 60
; :|||||
QY 32 QKRRNTIHEFKSAKTTLIKIDPALKIKTKVNTADQCANCRNKGLPFTCKAFVFDK 91
; :|||||
QY 61 ARKOCLWFPNSMSSGVKKEFGHEFDLYENKDYIRNCIIIGKGRSYKGTYSITKSGIKCOP 120
; :|||||
QY 92 ARKOCLWFPNSMSSGVKKEFGHEFDLYENKDYIRNCIIIGKGRSYKGTYSITKSGIKCOP 151
; :|||||
QY 121 WSSMIPHEHSFLPSSYRGKDLQENYCRNPRGEGGPMWCFSTSNPEVRYEYCDIPQCEVEEC 180
; :|||||
QY 152 WSSMIPHEHSFLPSSYRGKDLQENYCRNPRGEGGPMWCFSTSNPEVRYEYCDIPQCEVEEC 211
; :|||||
QY 181 MTCNGESYRGLMDHTESGKICORWDHOTPHRHKFLPERYDPKGFDDNYCRNPDGQPRPWC 240
; :|||||
QY 212 MTCNGESYRGLMDHTESGKICORWDHOTPHRHKFLPERYDPKGFDDNYCRNPDGQPRPWC 271
; :|||||
QY 241 YTLDPHTRWEYCAIKTCADNTMDTDVPLETTECIOGEGYRGTVNTIWNIGIPCORWDS 300
; :|||||
QY 272 YTLDPHTRWEYCAIKTCADNTMDTDVPLETTECIOGEGYRGTVNTIWNIGIPCORWDS 331
; :|||||
QY 301 QYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDMSHGQDC 360
; :|||||
QY 332 QYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDMSHGQDC 391
; :|||||
QY 361 YRGNGKNYMGNSLQTSRGLTCSMDKNMEDLHRHIFWEPDASKNLNENYCRNPDHDAHPW 420
; :|||||
QY 392 YRGNGKNYMGNSLQTSRGLTCSMDKNMEDLHRHIFWEPDASKNLNENYCRNPDHDAHPW 451
; :|||||
QY 421 CYTGNPLIPWDYCPISRCGDTTPTIV 447
; :|||||
QY 452 CYTGNPLIPWDYCPISRCGDTTPTIV 478
; :|||||

RESULT 13
US-08-605-221-2
; Sequence 2, Application US/08605221
; GENERAL INFORMATION:
; APPLICANT: NAKAMURA, Toshikazu
; TITLE OF INVENTION: THERAPEUTIC AGENT FOR DISORDER IN BRAIN AND NERVE
; FILE REFERENCE: 2520-0101P
; CURRENT APPLICATION NUMBER: US/08/605,221
; CURRENT FILING DATE: 1996-03-15
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 728
; TYPE: PRT
; ORGANISM: Homo sapiens
; OTHER INFORMATION: HGF-A (Full length HGF)
US-08-605-221-2

Query Match          99.9%; Score 2601; DB 10; Length 728;
Best Local Similarity 99.8%; Pred. No. 1.le-229;
Matches 446; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ERKRNTIHEFKSAKTTLIKIDPALKIKTKVNTADQCANCRNKGLPFTCKAFVFDK 60
; :|||||
```

0.

Search completed: April 25, 2003, 15:47:54  
Job time : 150 secs

## RESULT 15

03 00 330 333 I  
; Sequence 1, Application US/089309999

; APPLICANT: GEMBA, MUNEKAZU

; TITLE OF INVENTION: AGENT FO

FILE REFERENCE: 0050 1433 0 FCI  
; CURRENT APPLICATION NUMBER: US/08/930,999

EARLIER APPLICATION NUMBER: PCT/JP96/0106

EARLIER APPLICATION NUMBER: JP 96994/1995

**THIS PAGE BLANK (USPTO)**

Result No.	Query			DB	ID	Description
	Score	Match	Length			
1	2601	99.9	728	5	US-09-423-516B-2	Sequence 2, Appli
2	2601	99.9	728	5	US-09-423-516A-2	Sequence 2, Appli
3	2601	99.9	728	6	US-10-228-206-9	Sequence 9, Appli
4	2601	99.9	728	7	US-60-453-135-11153	Sequence 11153, A
5	2601	99.9	728	7	US-60-453-050-11153	Sequence 11153, A
6	2601	99.9	728	7	US-60-453-444-5842	Sequence 5842, Ap
7	2455	94.3	730	6	US-10-311-776-2	Sequence 2, Appli
8	2417.5	92.8	725	6	US-10-311-776-4	Sequence 4, Appli
9	1471	56.5	290	6	US-10-283-769-8	Sequence 8, Appli
10	1180.5	45.3	670	6	US-10-210-130-152	Sequence 152, App
11	1180.5	45.3	711	5	US-09-949-016-6981	Sequence 6981, Ap
12	1180.5	45.3	711	6	US-10-004-378A-118	Sequence 118, App
13	1180.5	45.3	711	6	US-10-004-378A-119	Sequence 119, App
14	1180.5	45.3	711	6	US-10-004-378A-120	Sequence 120, App
15	1168.5	44.9	712	6	US-10-210-130-142	Sequence 142, App
16	1164.5	44.7	543	5	US-09-724-676-63317	Sequence 63317, A
17	1164.5	44.7	543	5	US-09-724-676-63328	Sequence 63328, A
18	1164.5	44.7	543	5	US-09-724-676A-63317	Sequence 63317, A
19	1164.5	44.7	543	5	US-09-724-676A-63328	Sequence 63328, A
20	1164.5	44.7	568	5	US-09-724-676-63273	Sequence 63273, A
21	1164.5	44.7	568	5	US-09-724-676-63284	Sequence 63284, A
22	1164.5	44.7	568	5	US-09-724-676A-63273	Sequence 63273, A
23	1164.5	44.7	568	5	US-09-724-676A-63284	Sequence 63284, A
24	1164.5	44.7	674	5	US-09-724-676-63251	Sequence 63251, A
25	1164.5	44.7	674	5	US-09-724-676-63262	Sequence 63262, A
26	1164.5	44.7	674	5	US-09-724-676A-63251	Sequence 63251, A

QY 361 YRGNGKNYMGNLSTQSRGLTCSMDKNNMEDLHRHIFWEPDASKLNENYCRNPDDDAHGWP 420  
Db 392 YRGNGKNYMGNLSTQSRGLTCSMDKNNMEDLHRHIFWEPDASKLNENYCRNPDDDAHGWP 451  
QY 421 CYTGNPLIPWDYCPISRCGEGDTPPTIV 447  
Db 452 CYTGNPLIPWDYCPISRCGEGDTPPTIV 478

RESULT 2  
US-09-423-516A-2  
; Sequence 2, Application US/09423516A  
; GENERAL INFORMATION:  
; APPLICANT: Medical Research Council  
; TITLE OF INVENTION: HGF Polypeptides and their use in therapy  
; FILE REFERENCE: 1090-26  
; CURRENT APPLICATION NUMBER: US/09423516A  
; PRIOR FILING DATE: 2000-02-10  
; PRIOR APPLICATION NUMBER: PCT/GB98/01318  
; PRIOR FILING DATE: 1998-05-07  
; PRIOR APPLICATION NUMBER: GB 9709453.6  
; PRIOR FILING DATE: 1997-05-10  
; NUMBER OF SEQ ID NOS: 3  
; SOFTWARE: Patentin Ver. 2.0  
; SEQ ID NO: 2  
; LENGTH: 728  
; TYPE: PRT  
; ORGANISM: h. sapiens  
US-09-423-516A-2

Query Match 99.9%; Score 2601; DB 5; Length 728;  
Best Local Similarity 99.8%; Pred. No. 4.4e-206;  
Matches 446; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ERKRRNTIHEFKSAKTLTIKIDPALKIKTKVNTADQCANRTRNKGLPFTCKAFVFDK 60  
Db 32 QKRRNTIHEFKSAKTLTIKIDPALKIKTKVNTADQCANRTRNKGLPFTCKAFVFDK 91  
QY 61 ARKOCLEFPNSMSSGVKKEGHEFDLYENKDYIRNCIIIGKRSYKGTYSITKSGIKCOP 120  
Db 92 ARKOCLEFPNSMSSGVKKEGHEFDLYENKDYIRNCIIIGKRSYKGTYSITKSGIKCOP 151  
QY 121 WSSMIPHEHSLPSSYRGKDLQENYCRNPRGEGGPGWCFTSNPEVRYEVCIDIPOCSEVEC 180  
Db 152 WSSMIPHEHSLPSSYRGKDLQENYCRNPRGEGGPGWCFTSNPEVRYEVCIDIPOCSEVEC 211  
QY 181 MTCNGESYRGLMDHTESGKICQRDWHDQTPHHRKFLPERYDPKGFDDNNYCRNPDDDAHGWP 240  
Db 212 MTCNGESYRGLMDHTESGKICQRDWHDQTPHHRKFLPERYDPKGFDDNNYCRNPDDDAHGWP 271  
QY 241 YLDPHTRWECYCAIKTCADNTMNDTVPLETTICIQGEGYRGTVNTIWNIGIPCORWDS 300  
Db 272 YLDPHTRWECYCAIKTCADNTMNDTVPLETTICIQGEGYRGTVNTIWNIGIPCORWDS 331  
QY 301 QYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDMSHGQDC 360  
Db 332 QYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDMSHGQDC 391  
QY 361 YRGNGKNYMGNLSTQSRGLTCSMDKNNMEDLHRHIFWEPDASKLNENYCRNPDDDAHGWP 420  
Db 392 YRGNGKNYMGNLSTQSRGLTCSMDKNNMEDLHRHIFWEPDASKLNENYCRNPDDDAHGWP 451  
QY 421 CYTGNPLIPWDYCPISRCGEGDTPPTIV 447  
Db 452 CYTGNPLIPWDYCPISRCGEGDTPPTIV 478

RESULT 3  
US-10-228-206-9  
; Sequence 9, Application US/10228206  
; GENERAL INFORMATION:  
; APPLICANT: Hitachi LTD.  
; TITLE OF INVENTION: HGF-INTERACTING PROTEINS AND USE THEREOF

FILE REFERENCE: H0100932A1  
; CURRENT APPLICATION NUMBER: US/10/228,206  
; CURRENT FILING DATE: 2002-08-27  
; PRIOR APPLICATION NUMBER: US 60/326,423  
; PRIOR FILING DATE: 2001-10-03  
; NUMBER OF SEQ ID NOS: 15  
; SOFTWARE: Patentin version 3.1  
; SEQ ID NO: 9  
; LENGTH: 728  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-228-206-9

Query Match 99.9%; Score 2601; DB 6; Length 728;  
Best Local Similarity 99.8%; Pred. No. 4.4e-206;  
Matches 446; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ERKRRNTIHEFKSAKTLTIKIDPALKIKTKVNTADQCANRTRNKGLPFTCKAFVFDK 60  
Db 32 QKRRNTIHEFKSAKTLTIKIDPALKIKTKVNTADQCANRTRNKGLPFTCKAFVFDK 91  
QY 61 ARKOCLEFPNSMSSGVKKEGHEFDLYENKDYIRNCIIIGKRSYKGTYSITKSGIKCOP 120  
Db 92 ARKOCLEFPNSMSSGVKKEGHEFDLYENKDYIRNCIIIGKRSYKGTYSITKSGIKCOP 151  
QY 121 WSSMIPHEHSLPSSYRGKDLQENYCRNPRGEGGPGWCFTSNPEVRYEVCIDIPOCSEVEC 180  
Db 152 WSSMIPHEHSLPSSYRGKDLQENYCRNPRGEGGPGWCFTSNPEVRYEVCIDIPOCSEVEC 211  
QY 181 MTCNGESYRGLMDHTESGKICQRDWHDQTPHHRKFLPERYDPKGFDDNNYCRNPDDDAHGWP 240  
Db 212 MTCNGESYRGLMDHTESGKICQRDWHDQTPHHRKFLPERYDPKGFDDNNYCRNPDDDAHGWP 271  
QY 241 YLDPHTRWECYCAIKTCADNTMNDTVPLETTICIQGEGYRGTVNTIWNIGIPCORWDS 300  
Db 272 YLDPHTRWECYCAIKTCADNTMNDTVPLETTICIQGEGYRGTVNTIWNIGIPCORWDS 331  
QY 301 QYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDMSHGQDC 360  
Db 332 QYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDMSHGQDC 391  
QY 361 YRGNGKNYMGNLSTQSRGLTCSMDKNNMEDLHRHIFWEPDASKLNENYCRNPDDDAHGWP 420  
Db 392 YRGNGKNYMGNLSTQSRGLTCSMDKNNMEDLHRHIFWEPDASKLNENYCRNPDDDAHGWP 451  
QY 421 CYTGNPLIPWDYCPISRCGEGDTPPTIV 447  
Db 452 CYTGNPLIPWDYCPISRCGEGDTPPTIV 478

RESULT 4  
US-60-453-135-11153  
; Sequence 11153, Application US/60453135  
; GENERAL INFORMATION:  
; APPLICANT: CARGILL, Michele  
; APPLICANT: IAKOUBOVA, Olga  
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH  
; TITLE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CLO01456  
; CURRENT APPLICATION NUMBER: US/60/453,135  
; CURRENT FILING DATE: 2003-03-10  
; NUMBER OF SEQ ID NOS: 82762  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO: 11153  
; LENGTH: 728  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-60-453-135-11153

Query Match 99.9%; Score 2601; DB 7; Length 728;  
Best Local Similarity 99.8%; Pred. No. 4.4e-206;  
Matches 446; Conservative 1; Mismatches 0; Indels 0; Gaps 0;



```
Qy 1 ERKRNTTIEFKSAKTTLIKIDPALKIKTKKVNADOCANRCKTRNKGLPFTCKAFVFDK 60
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 32 QRKRNTTIEFKSAKTTLIKIDPALKIKTKKVNADOCANRCKTRNKGLPFTCKAFVFDK 91
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Qy 61 ARKQCLWFPFNSMSGVKKKEGHEFDLYENKDYIRNCIIGKGRSYKGTVSTKSGIKQCP 120
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 92 ARKQCLWFPFNSMSGVKKKEGHEFDLYENKDYIRNCIIGKGRSYKGTVSTKSGIKQCP 151
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Qy 121 WSSMIPHEHSFLPSSYRGKDLQENYCRNPRGEGGPGWCFSTNPEVRYEVCIDIPOCSEVEC 180
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 152 WSSMIPHEHSFLPSSYRGKDLQENYCRNPRGEGGPGWCFSTNPEVRYEVCIDIPOCSEVEC 211
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Qy 181 MTCNGESYRGLMDHTESGKICQWRDHQTPHRRKFLPERYPDKGFDNDYCRNPDGQPRWC 240
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 212 MTCNGESYRGLMDHTESGKICQWRDHQTPHRRKFLPERYPDKGFDNDYCRNPDGQPRWC 271
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Qy 241 YLDPHTRWEYCAIKTCADNTMNDTVPLETECIGOGEGYRGTVNTIWNIPCORWDS 300
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 272 YLDPHTRWEYCAIKTCADNTMNDTVPLETECIGOGEGYRGTVNTIWNIPCORWDS 331
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Qy 301 QYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSOIPNCDMSHGDC 360
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 332 QYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSOIPNCDMSHGDC 391
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Qy 361 YRGNGKNYMGNLSTQSRGLTCSMDKKNMEDLHRHIFWEPDASKLKNENYCRNPDDDAHGPW 420
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 392 YRGNGKNYMGNLSTQSRGLTCSMDKKNMEDLHRHIFWEPDASKLKNENYCRNPDDDAHGPW 451
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Qy 421 CYTGNPLIPWDYCPISRCGEGDTTPTIV 447
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 452 CYTGNPLIPWDYCPISRCGEGDTTPTIV 478
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
RESULT 5
US-60-453-050-11153
; Sequence 11153, Application US/60453050
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; TITLE OF INVENTION: STENOSIS, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001457
; CURRENT APPLICATION NUMBER: US/60/453.050
; CURRENT FILING DATE: 2003-03-10
; NUMBER OF SEQ ID NOS: 82762
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11153
; LENGTH: 728
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-453-050-11153
```

```
Query Match 99.9%; Score 2601; DB 7; Length 728;
Best Local Similarity 99.8%; Pred. No. 4.4e-206;
Matches 446; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ERKRNTTIEFKSAKTTLIKIDPALKIKTKKVNADOCANRCKTRNKGLPFTCKAFVFDK 60
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 32 QRKRNTTIEFKSAKTTLIKIDPALKIKTKKVNADOCANRCKTRNKGLPFTCKAFVFDK 91
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Qy 61 ARKQCLWFPFNSMSGVKKKEGHEFDLYENKDYIRNCIIGKGRSYKGTVSTKSGIKQCP 120
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 92 ARKQCLWFPFNSMSGVKKKEGHEFDLYENKDYIRNCIIGKGRSYKGTVSTKSGIKQCP 151
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Qy 121 WSSMIPHEHSFLPSSYRGKDLQENYCRNPRGEGGPGWCFSTNPEVRYEVCIDIPOCSEVEC 180
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 152 WSSMIPHEHSFLPSSYRGKDLQENYCRNPRGEGGPGWCFSTNPEVRYEVCIDIPOCSEVEC 211
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Qy 181 MTCNGESYRGLMDHTESGKICQWRDHQTPHRRKFLPERYPDKGFDNDYCRNPDGQPRWC 240
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 212 MTCNGESYRGLMDHTESGKICQWRDHQTPHRRKFLPERYPDKGFDNDYCRNPDGQPRWC 271
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Qy 241 YLDPHTRWEYCAIKTCADNTMNDTVPLETECIGOGEGYRGTVNTIWNIPCORWDS 300
```

```
Db 272 YLDPHTRWEYCAIKTCADNTMNDTVPLETECIGOGEGYRGTVNTIWNIPCORWDS 331
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Qy 301 QYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSOIPNCDMSHGDC 360
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 332 QYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSOIPNCDMSHGDC 391
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Qy 361 YRGNGKNYMGNLSTQSRGLTCSMDKKNMEDLHRHIFWEPDASKLKNENYCRNPDDDAHGPW 420
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 392 YRGNGKNYMGNLSTQSRGLTCSMDKKNMEDLHRHIFWEPDASKLKNENYCRNPDDDAHGPW 451
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Qy 421 CYTGNPLIPWDYCPISRCGEGDTTPTIV 447
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 452 CYTGNPLIPWDYCPISRCGEGDTTPTIV 478
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
RESULT 6
US-60-455-444-5842
; Sequence 5842, Application US/60455444
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: BEGOVICH, Ann
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: RHEUMATOID ARTHRITIS, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001455
; CURRENT APPLICATION NUMBER: US/60/455.444
; CURRENT FILING DATE: 2003-03-18
; NUMBER OF SEQ ID NOS: 50986
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5842
; LENGTH: 728
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-455-444-5842
```

```
Query Match 99.9%; Score 2601; DB 7; Length 728;
Best Local Similarity 99.8%; Pred. No. 4.4e-206;
Matches 446; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ERKRNTTIEFKSAKTTLIKIDPALKIKTKKVNADOCANRCKTRNKGLPFTCKAFVFDK 60
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 32 QRKRNTTIEFKSAKTTLIKIDPALKIKTKKVNADOCANRCKTRNKGLPFTCKAFVFDK 91
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Qy 61 ARKQCLWFPFNSMSGVKKKEGHEFDLYENKDYIRNCIIGKGRSYKGTVSTKSGIKQCP 120
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 92 ARKQCLWFPFNSMSGVKKKEGHEFDLYENKDYIRNCIIGKGRSYKGTVSTKSGIKQCP 151
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Qy 121 WSSMIPHEHSFLPSSYRGKDLQENYCRNPRGEGGPGWCFSTNPEVRYEVCIDIPOCSEVEC 180
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 152 WSSMIPHEHSFLPSSYRGKDLQENYCRNPRGEGGPGWCFSTNPEVRYEVCIDIPOCSEVEC 211
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Qy 181 MTCNGESYRGLMDHTESGKICQWRDHQTPHRRKFLPERYPDKGFDNDYCRNPDGQPRWC 240
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 212 MTCNGESYRGLMDHTESGKICQWRDHQTPHRRKFLPERYPDKGFDNDYCRNPDGQPRWC 271
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Qy 241 YLDPHTRWEYCAIKTCADNTMNDTVPLETECIGOGEGYRGTVNTIWNIPCORWDS 300
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 272 YLDPHTRWEYCAIKTCADNTMNDTVPLETECIGOGEGYRGTVNTIWNIPCORWDS 331
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Qy 301 QYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSOIPNCDMSHGDC 360
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 332 QYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSOIPNCDMSHGDC 391
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Qy 361 YRGNGKNYMGNLSTQSRGLTCSMDKKNMEDLHRHIFWEPDASKLKNENYCRNPDDDAHGPW 420
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 392 YRGNGKNYMGNLSTQSRGLTCSMDKKNMEDLHRHIFWEPDASKLKNENYCRNPDDDAHGPW 451
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Qy 421 CYTGNPLIPWDYCPISRCGEGDTTPTIV 447
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 452 CYTGNPLIPWDYCPISRCGEGDTTPTIV 478
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
RESULT 7
```

```
US-10-311-776-2
; Sequence 2, Application US/10311776
; GENERAL INFORMATION:
; APPLICANT: MIYAKE, Masashi
; APPLICANT: IWABUCHI, Shigehiro
; APPLICANT: SUZUTA, Yasuyuki
; TITLE OF INVENTION: Canine Hepatocyte Growth Factor
; FILE REFERENCE: 522.1023
; CURRENT APPLICATION NUMBER: US/10/311.776
; PRIOR FILING DATE: 2002-12-18
; PRIOR APPLICATION NUMBER: JP2000-187724
; PRIOR FILING DATE: 2000-06-22
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 2
; LENGTH: 730
; TYPE: PRT
; ORGANISM: Canis familiaris
; FEATURE:
; NAME/KEY: Variation
; LOCATION: (138)
; OTHER INFORMATION: Xaa = Gly
US-10-311-776-2

Query Match          94.3%; Score 2455; DB 6; Length 730;
Best Local Similarity 92.8%; Pred. No. 4.7e-194;
Matches 415; Conservative 19; Mismatches 13; Indels 0; Gaps 0;

QY 1 ERKRRNTIHEFKSAKTLIKIDPALKIKTKVNTADQCANRCTRNKGLPFTCKAFVDEK 60
Db 32 QKRRNTLHEFKSAKTLIKEDPDLKIKTKMNTADQCANCTRNKGLPFTCKAFVDEK 91
QY 61 ARKQCLWFPNSMSSGVKKEGHEFDLYENKDYIRNCIIIGKRSYKGTVSITKSGIKCOP 120
Db 92 ARKCLWFPNSMTSGYKKEGHEFDLYENKDYIRNCIIIGKRSYKGTVSITKSGIKCOP 151
QY 121 WSMIPHEHSLPSSYRGKDLQENYCRNPRGEGGPGWCFTSNPEVRYEYVCDIPQSEVEK 180
Db 152 WNSMIPHEHSLPSSYRGKDLQENYCRNPRGEGGPGWCFTSNPEVRYEYVCDIPQSEVEK 211
QY 181 MTCNGESYRGLMDHTESGKICQWDHQTPHRHKFLPERYPDKGFDNNYCRNPDGKPRPWC 240
Db 212 MTCNGESYRGPMDHTESGKICQWDHQTPHRHKFLPERYPDKGFDNNYCRNPDGKPRPWC 271
QY 241 YTLDPHTWEYCAIKTCADNTMDTVPLETTTCIOGEGYRGTVNTIWNIGPCQWRDS 300
Db 272 YTLDPHTWEYCAIKTCADNTMDTVPLETTTCIOGEGYRGTVNTIWNIGPCQWRDS 331
QY 301 QYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDMSHGQDC 360
Db 332 QYPHODHTPENFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDMSHGQDC 391
QY 361 YRGNGKNYMGLNSTRSGLTCSMWKKNMEDLHRHIFWEPDASKLNENYCRNPDGDAHGWP 420
Db 392 YRGNGKNYMGLNSTRSGLTCSMWKKNMEDLHRHIFWEPDASKLNENYCRNPDGDAHGWP 451
QY 421 CYTGNPLIPWDYCPISRCGEGDTTPIV 447
Db 452 CYTGNPLIPWDYCPISRCGEGDTTPIV 478

RESULT 8
US-10-311-776-4
; Sequence 4, Application US/10311776
; GENERAL INFORMATION:
; APPLICANT: MIYAKE, Masashi
; APPLICANT: IWABUCHI, Shigehiro
; APPLICANT: SUZUTA, Yasuyuki
; TITLE OF INVENTION: Canine Hepatocyte Growth Factor
; FILE REFERENCE: 522.1023
; CURRENT APPLICATION NUMBER: US/10/311.776
; CURRENT FILING DATE: 2002-12-18
; PRIOR APPLICATION NUMBER: JP2000-187724

US-10-283-769-8
; Sequence 8, Application US/10283769
; GENERAL INFORMATION:
; APPLICANT: Chan, Andrew ML
; APPLICANT: Rubin, Jeffrey L.
; APPLICANT: Bottaro, Donald P.
; APPLICANT: Aaronson, Stuart A.
; APPLICANT: Stahl, Stephen J.
; APPLICANT: Wingfield, Paul T.
; APPLICANT: Cioce, Vittoria
; TITLE OF INVENTION: TRUNCATED HEPATOCYTE GROWTH FACTOR
; FILE REFERENCE: 14014.0269
; CURRENT APPLICATION NUMBER: US/10/283.769
; PRIOR FILING DATE: 2002-10-30
; PRIOR APPLICATION NUMBER: US/08/484.841
; PRIOR FILING DATE: 2002-03-13
; PRIOR APPLICATION NUMBER: 08/130.134
; PRIOR FILING DATE: 1993-10-04
; PRIOR APPLICATION NUMBER: 07/655.502
; PRIOR FILING DATE: 1991-02-15
; PRIOR APPLICATION NUMBER: 07/582.063

US-10-283-769-8
; Sequence 8, Application US/10283769
; GENERAL INFORMATION:
; APPLICANT: Chan, Andrew ML
; APPLICANT: Rubin, Jeffrey L.
; APPLICANT: Bottaro, Donald P.
; APPLICANT: Aaronson, Stuart A.
; APPLICANT: Stahl, Stephen J.
; APPLICANT: Wingfield, Paul T.
; APPLICANT: Cioce, Vittoria
; TITLE OF INVENTION: TRUNCATED HEPATOCYTE GROWTH FACTOR
; FILE REFERENCE: 14014.0269
; CURRENT APPLICATION NUMBER: US/10/283.769
; PRIOR FILING DATE: 2002-10-30
; PRIOR APPLICATION NUMBER: US/08/484.841
; PRIOR FILING DATE: 2002-03-13
; PRIOR APPLICATION NUMBER: 08/130.134
; PRIOR FILING DATE: 1993-10-04
; PRIOR APPLICATION NUMBER: 07/655.502
; PRIOR FILING DATE: 1991-02-15
; PRIOR APPLICATION NUMBER: 07/582.063
```

```
; PRIOR FILING DATE: 1990-09-14
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 290
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Artificial Sequence :/ Note = synthetic construct
US-10-283-769-8

Query Match          56.5%; Score 1471; DB 6; Length 290;
Best Local Similarity 99.6%; Pred. No. 2.3e-113;
Matches 256; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1  ERKRRNTHEFKSAKTLIKIDPALKIKTKKVTADQCANCTRNKGLPFTCKAFVFDK 60
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 32  QRKRRNTHEFKSAKTLIKIDPALKIKTKKVTADQCANCTRNKGLPFTCKAFVFDK 91

QY 61  ARKQCLWFFPNSMSGVKKFEGHEFDLYENKDYIRNCIIIGKGRSYKGTVSIITKSGIKCOP 120
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 92  ARKQCLWFFPNSMSGVKKFEGHEFDLYENKDYIRNCIIIGKGRSYKGTVSIITKSGIKCOP 151

QY 121  WSSMIPHEHSFLPSSYRGKDLQENYCRNPRGEGGWCFTSNPEVRYEYCDIPIQCSVEVC 180
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 152  WSSMIPHEHSFLPSSYRGKDLQENYCRNPRGEGGWCFTSNPEVRYEYCDIPIQCSVEVC 211

QY 181  MTCNGESYGLMDHTESGKICQWRDQHTPHRHKFLPERYPDKGFDNNYCRNPDGQPRPWC 240
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 212  MTCNGESYGLMDHTESGKICQWRDQHTPHRHKFLPERYPDKGFDNNYCRNPDGQPRPWC 271

QY 241  YTLDPHTRWEYCAIKTC 257
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 272  YTLDPHTRWEYCAIKTC 288

RESULT 10
US-10-210-130-152
; Sequence 152, Application US/10210130
; GENERAL INFORMATION:
; APPLICANT: Zerhusen, Bryan D.
; APPLICANT: Patturajan, Meera
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Miller, Charles E.
; APPLICANT: Rieger, Daniel K.
; APPLICANT: Pena, Carol E.A.
; APPLICANT: Shimkets, Richard A.
; APPLICANT: Li, Li
; APPLICANT: Berghs, Constance
; APPLICANT: Zhong, Mei
; APPLICANT: Casman, Stacie J.
; APPLICANT: Voss, Edward Z.
; APPLICANT: Boldog, Ferenc L.
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Smithson, Glennda
; APPLICANT: Ji, Weizhen
; APPLICANT: Gorman, Linda
; APPLICANT: Vernet, Corine A.M.
; APPLICANT: Leite, Mario W.
; APPLICANT: Guo, Xiaojia Sasha
; APPLICANT: Anderson, David W.
; APPLICANT: Spytek, Kimberly A.
; APPLICANT: Gerlach, Valerie
; APPLICANT: Burgess, Catherine E.
; APPLICANT: Khrantsov, Nikolai V.
; APPLICANT: Ort, Tatiana
; APPLICANT: Ellerman, Karen
; APPLICANT: Rastelli, Luca
; APPLICANT: Agee, Michele L.
; APPLICANT: Chaudhuri, Amitabha
; APPLICANT: Chant, John S.
; APPLICANT: DiPippo, Vincent A.
; APPLICANT: Edinger, Shlomit R.

; APPLICANT: Eisen, Andrew J.
; APPLICANT: Gangolli, Esha A.
; APPLICANT: Giot, Loic
; APPLICANT: Ooi, Chean Eng
; APPLICANT: Rothenberg, Mark E.
; APPLICANT: Spaderna, Steven K.
; APPLICANT: Hjalt, Tord
; APPLICANT: Liu, Xiaohong
; APPLICANT: Taupier, Raymond J., Jr.
; APPLICANT: Catterton, Elina
; APPLICANT: Shenoy, Suresh G.
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 21402-416C (Cura-716 SMT)
; CURRENT APPLICATION NUMBER: US/10/210,130
; CURRENT FILING DATE: 2002-08-01
; PRIOR APPLICATION NUMBER: 60/309,501
; PRIOR FILING DATE: 2001-08-02
; PRIOR APPLICATION NUMBER: 60/316,508
; PRIOR FILING DATE: 2001-08-31
; PRIOR APPLICATION NUMBER: 60/354,655
; PRIOR FILING DATE: 2002-02-05
; PRIOR APPLICATION NUMBER: 60/310,291
; PRIOR FILING DATE: 2001-08-03
; PRIOR APPLICATION NUMBER: 60/383,887
; PRIOR FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: 60/310,951
; PRIOR FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: 60/323,936
; PRIOR FILING DATE: 2001-09-21
; PRIOR APPLICATION NUMBER: 60/381,039
; PRIOR FILING DATE: 2002-05-16
; PRIOR APPLICATION NUMBER: 60/311,292
; PRIOR FILING DATE: 2001-08-09
; PRIOR APPLICATION NUMBER: 60/311,979
; PRIOR FILING DATE: 2001-08-13
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 369
; SOFTWARE: CuraSeqList version 0.1
; SEQ ID NO 152
; LENGTH: 670
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-210-130-152

Query Match          45.3%; Score 1180.5; DB 6; Length 670;
Best Local Similarity 47.4%; Pred. No. 4.6e-89;
Matches 197; Conservative 62; Mismatches 148; Indels 9; Gaps 4;

QY 33  VNTADQCANRCTRNKGLPFTCTKAFVFDKARKOCLMPPFNMSMSGVKKFEGHEFDLYENKD 92
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 64  VADABECAGRC---GPLMDCAFRHYNVSSHGCOLLPWTQHSPTLRRLRRSGRCDLFQKKD 119

QY 93  YIRNCIIIGKGRSYKGTVSIITKSGIKCQWPSSMTIPHEHSFLPSSYRGKDLQENYCRNPRGE 152
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 120  YVRTCIMNNGVGYRGTMATTVGGLPCQASHKFPNDKHKTPTLRNG--LEENFCRNPDDG 177

QY 153  EGGPWCFTSNPEVRYEYCDIPIQCSVEVCMTCSNGESYGLMDHTESGKICQWRDQHTPHRH 212
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 178  PGGPWCYTTDPVAFQSCGKISCREAACVWCNGEYRGAVDRATESGRCQWRDQLQHPHQH 237

QY 213  KFLPERYPDKGFDNNYCRNPDGQPRPWCYTLDPHTRWECYCAIKTCADNTMNDTDPLETT 272
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 238  PFEPGKFLDQGLDDNYCRNPDGSERPWCYTTDPIEREFCDLPRCSEAQPROEA--TTV 295

QY 273  ECIOGEGYRGTVNTIWNIGIPCORWDSOYPHEHDMTTPENFKCKDLRENYCRNPDGSESP 332
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 296  SCFRGKEGYRGTAHTTAGVPCQWRDQAIHQHRTTPEKYACKDLRENYCRNPDGSEAP 355

QY 333  WCFTTDPNIRVGYCSQIPNC--DMSHGQDCYRGNGKNMGNLSQTRSGLTCSWMDKNMEDL 391
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 356  WCFTLRPGMRAAFYQIRRCTDDVDPQDCYHGAGEQYRGTVSKTRKGVQCORWSAETPHK 415

QY 392  HRHIFWEPDASKLINENYCRNPDGDDAHGPWCYTGNNPLIPWDYCPISRCGEGTPTTV 447
```

[illegible]

```
Db 282 SCFRGKGEYRGNTAGVPCQWDAQIPHQHRTPEKYACKDLRENFRCNPDGSEAP 341
QY 333 WCFTTDPNIRVGYCSQIPNC--DMSHGDCYRGNGKNTMGNLSOTRSGLTCSMWKDNMEDL 391
Db 342 WCFTLRGCMRAAFYQYLRCTDDVRRPQDCYHGAGEQYRGVSKTRKGVCQORWSAETPHK 401
QY 392 HRHIFWEPDASKLNNENYCRNPDHDAHGWCYTGPNLPIDWYCPISRCGDDTPTIV 447
Db 402 PQFTFTSEPHAQLEENFCRNPDGSHGWPWCYTMDPRTPTFDYCALRRCADDDPPSIL 457

RESULT 13
US-10-004-378A-119
; Sequence 119, Application US/10004378A
; GENERAL INFORMATION:
; APPLICANT: Li, Li
; APPLICANT: Furtak, Kazarzyna
; APPLICANT: Perna, Amanda
; APPLICANT: Patturajan, Meera
; APPLICANT: Shimkets, Richard A
; APPLICANT: Guo, Xiaojia Sasha
; APPLICANT: Casman, Stacie J
; APPLICANT: Burgess, Catherine E
; APPLICANT: Malyankar, Uriel M
; APPLICANT: Tchernev, Velizar T
; APPLICANT: Vernet, Corrine A
; APPLICANT: Spytek, Kimberly A
; APPLICANT: Agee, Michele
; APPLICANT: Rastelli, Luca
; APPLICANT: Shenoy, Suresh G
; APPLICANT: Grosse, William M
; APPLICANT: Alsobrook II, John P
; APPLICANT: Lepley, Denise M
; APPLICANT: Gerlach, Valerie
; APPLICANT: Edinger, Schlomit
; APPLICANT: MacDougall, John R
; APPLICANT: Peyman, John A
; APPLICANT: Gunther, Erik
; APPLICANT: Stone, David J
; APPLICANT: Ellerman, Karen
; APPLICANT: Gangolli, Esha A
; TITLE OF INVENTION: Novel Human Proteins, Polynucleotides Encoding Them and
; FILE REFERENCE: 21402-179
; CURRENT APPLICATION NUMBER: US/10/004.378A
; CURRENT FILING DATE: 2001-10-24
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: 60/242,882
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: 60/242,765
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: 60/300,206
; PRIOR FILING DATE: 2001-06-22
; PRIOR APPLICATION NUMBER: 60/242,789
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: 60/242,768
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: 60/242,767
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: 60/243,622
; PRIOR FILING DATE: 2000-10-26
; PRIOR APPLICATION NUMBER: 60/273,047
; PRIOR FILING DATE: 2001-03-02
; PRIOR APPLICATION NUMBER: 60/243,591
; PRIOR FILING DATE: 2000-10-26
; PRIOR APPLICATION NUMBER: 60/243,950
; PRIOR FILING DATE: 2000-10-27
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 191
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 119
; LENGTH: 711
; TYPE: PRT
```

```
; ORGANISM: Homo sapiens
US-10-004-378A-119
Query Match 45.3%; Score 1180.5; DB 6; Length 711;
Best Local Similarity 47.4%; Pred. No. 4.9e-89;
Matches 197; Conservative 62; Mismatches 148; Indels 9; Gaps 4;

QY 33 VNTADOCANRCRTRNKGLPFTCKAFVFDKARKQCLWFPENSMSSGVKKFEGFHEFDLYENKD 92
Db 50 VADAEACAGRC---GPLMDCRAFYHNVSSHGCCQLLPWQHSPHTRLRRSGRCDLFQKDK 105
QY 93 YIRNCIIIGKGRSYKGTVSIITKSGIKCQWPSSMIPEHSIFLPSYSYRGKDLQNYCNRPRGE 152
Db 106 YVRTCIMNNGVGYRGTMATTVGGLPQCAWSKFFNDHKYTPTLRNG--LEENFCRNPDG 163
QY 153 EGGPWCFTSNPEVRYEVCIDIPOCSECEMTCNGESYRGLMDHTESGKICORWDHQTPHRH 212
Db 164 PGGWCYTTDDPAVRFQSGGKISCREAACVWCNGEYRGAVDRTESGRCQQRWDLQHPHOH 223
QY 213 KFLPERYPDKGFDNNYCRNPDGQPRPWCYTTLDPHTRWYECYCAIKTCADNTMDTDVPLETT 272
Db 224 PFEFGKFLDQGLDDNYCRNPDGSEPRWCYTTPQIEREFCDLPRCGSEAQPRQEA--TTV 281
QY 273 ECIQGQGEYRGTVNTIWNIGIPCQRWDSQYPHEHDMTPPENFKCKDLRENYCRNPDGSESP 332
Db 282 SCFRGKGEYRGNTAGVPCQWDAQIPHQHRTPEKYACKDLRENFRCNPDGSEAP 341
QY 333 WCFTTDPNIRVGYCSQIPNC--DMSHGDCYRGNGKNTMGNLSOTRSGLTCSMWKDNMEDL 391
Db 342 WCFTLRGCMRAAFYQYLRCTDDVRRPQDCYHGAGEQYRGVSKTRKGVCQORWSAETPHK 401
QY 392 HRHIFWEPDASKLNNENYCRNPDHDAHGWCYTGPNLPIDWYCPISRCGDDTPTIV 447
Db 402 PQFTFTSEPHAQLEENFCRNPDGSHGWPWCYTMDPRTPTFDYCALRRCADDDPPSIL 457

RESULT 14
US-10-004-378A-120
; Sequence 120, Application US/10004378A
; GENERAL INFORMATION:
; APPLICANT: Li, Li
; APPLICANT: Furtak, Kazarzyna
; APPLICANT: Perna, Amanda
; APPLICANT: Patturajan, Meera
; APPLICANT: Shimkets, Richard A
; APPLICANT: Guo, Xiaojia Sasha
; APPLICANT: Casman, Stacie J
; APPLICANT: Burgess, Catherine E
; APPLICANT: Malyankar, Uriel M
; APPLICANT: Tchernev, Velizar T
; APPLICANT: Vernet, Corrine A
; APPLICANT: Spytek, Kimberly A
; APPLICANT: Agee, Michele
; APPLICANT: Rastelli, Luca
; APPLICANT: Shenoy, Suresh G
; APPLICANT: Grosse, William M
; APPLICANT: Alsobrook II, John P
; APPLICANT: Lepley, Denise M
; APPLICANT: Gerlach, Valerie
; APPLICANT: Edinger, Schlomit
; APPLICANT: MacDougall, John R
; APPLICANT: Peyman, John A
; APPLICANT: Gunther, Erik
; APPLICANT: Stone, David J
; APPLICANT: Ellerman, Karen
; APPLICANT: Gangolli, Esha A
; TITLE OF INVENTION: Novel Human Proteins, Polynucleotides Encoding Them and
; FILE REFERENCE: 21402-179
; CURRENT APPLICATION NUMBER: US/10/004.378A
; CURRENT FILING DATE: 2001-10-24
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: 60/242,882
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: 60/242,765
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: 60/300,206
; PRIOR FILING DATE: 2001-06-22
; PRIOR APPLICATION NUMBER: 60/242,789
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: 60/242,768
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: 60/242,767
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: 60/243,622
; PRIOR FILING DATE: 2000-10-26
; PRIOR APPLICATION NUMBER: 60/273,047
; PRIOR FILING DATE: 2001-03-02
; PRIOR APPLICATION NUMBER: 60/243,591
; PRIOR FILING DATE: 2000-10-26
; PRIOR APPLICATION NUMBER: 60/243,950
; PRIOR FILING DATE: 2000-10-27
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 191
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 119
; LENGTH: 711
; TYPE: PRT
```

QV 4 ERNTTIEEKYSAKTTLIKIDPALKIK--TKKVNTADOCANRCTRNKGLPFTCKAFVFDKA 61  
Matches 206; Conservative 73; Mismatches 154; Indels 17; Gaps 8

Db 19 QRSPLNDFEVLRGTELQLOAVVPGMQEDVADAEBCAGRC---GPLMDCRAFYHVS 74  
Qy 62 RKQCLWFPFNSMGVKKKEFGHEFDLYENKDYIRNCIIGKGRSYKGTYSITKSGIKCOPW 121  
Db 75 SHGCOLLFWTQHSPTLRHSHGRCDLQFEKDYIRTCINNGVGYRGWTWATTVGLSCQAW 134  
Qy 122 SSMIPHEHSFLPSSYRGKDLQENYCRNPRGEGGPGWCFTSNPEVRYEYCDIPQCSEVECM 181  
Db 135 SHKFPNDHRYMPTLRNG--LEENFCRNPDPGGPGWCHTTPAVRFQSCGKISCRSAACV 192  
Qy 182 TCNGESYRGLMDHTESGKICORWDHQTPHRHKFLPERYPDKGFDNDNYCRNPDGQPRWCY 241  
Db 193 WCNGEYRGAVDRTESGRECORWDLQHPHQPFBPGKYPDQGLDDNYCRNPDGGERPWCY 252  
Qy 242 TLDPHTRWEYCAIKTCAADNTMNDTVDPLETTECIQOGGEGYRGVTNTIWNIGIPCORWDSQ 301  
Db 253 TTDQIEREFCDLPRCSQAQPROEA--TSVSCFRGKGEYRGTTANTTTAGVPCORWDAQ 310  
Qy 302 YPHEHMDTPENFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNC-DMSHGODC 360  
Db 311 IPHQHRETPERYACKDLRENEFCRNPDPGSEAPWCFTLRPGMRVGFYQIRRCTDDVVRPOGC 370  
Qy 361 YRGCKNYMGNLSQTRSLTCSMWDKNMEDLHRHFW---EPDASKLNEYCRNPDDDAH 417  
Db 371 YHGAGEQYRGTVSKTRKGVQCOR--ASAETHKPOFTTSEPHA-OLEENFCRDPDGDY 427  
Qy 418 GPWCYTGNPLIPWDYCPISRCGEGTTPITIV 447  
Db 428 GPWCYTMDPRTPFDYCALRCADDOPPSIL 457

Search completed: April 25, 2003, 15:48:37  
Job time : 41 secs

**THIS PAGE BLANK (uspto)**



GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 25, 2003, 15:48:43 ; Search time 46 Seconds  
(without alignments)  
934.176 Million cell updates/sec

Title: US-09-674-377B-1

Perfect score: 2604

Sequence: 1 ERRRNTIHEFKSAKTLI.....IPWDYCPISRCGDTPTIV 447

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR\_73.\*  
1: piri:.\*  
2: piri2:.\*  
3: piri3:.\*  
4: piri4:.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2601	99.9	728	1 JH0579	hepatocyte growth
2	2408	92.5	728	1 A60185	hepatocyte growth
3	2401	92.2	728	1 A35644	hepatocyte growth
4	1864.5	71.6	710	1 I51283	hepatocyte growth
5	1737.5	66.7	411	2 I51285	hepatocyte growth
6	1180.5	45.3	711	1 A47136	macrophage-stimula
7	1144	43.9	716	1 JC5061	macrophage-stimula
8	1125	43.2	716	1 A40332	macrophage-stimula
9	1066	40.9	790	1 PLPG	plasma
10	1059	40.7	812	1 PLMS	plasma
11	1053.5	40.5	810	2 I46260	plasma
12	1044.5	40.1	810	2 B30848	plasma
13	1042	40.0	810	1 PLBU	plasma
14	1038	39.9	812	1 PLBO	plasma
15	943.5	36.2	2869	2 T18518	apolipoprotein(a)
16	879	33.8	4548	1 S00657	apolipoprotein(a)
17	843.5	32.4	1420	2 A32869	apolipoprotein(a)
18	457	17.5	460	2 B61345	plasma
19	454	17.4	455	2 A61545	plasma
20	408.5	15.7	336	2 S33879	plasma
21	357.5	13.7	625	1 TBBO	thrombin
22	353	13.6	169	2 A40522	plasma
23	352.5	13.5	618	2 A35827	thrombin
24	350	13.4	622	1 TBHU	thrombin
25	344	13.2	617	2 S10511	thrombin
26	307.5	11.8	559	1 A35029	t-plasminogen acti
27	303	11.6	562	1 UKHUT	t-plasminogen acti
28	297.5	11.4	559	1 A29941	t-plasminogen acti
29	270	10.4	120	2 B61545	plasma

#### RESULT 1

JH0579

hepatocyte growth factor precursor [validated] - human

N:Alternate names: hepaoleitin A; scatter factor

C:Species: Homo sapiens (man)

C>Date: 17-Aug-1992 #sequence-revision 17-Aug-1992 #text-change 08-Dec-2000

C:Accession: JH0579; JH0333; A41140; B36677; A33512; A39006; PH0114; A37796;

R:Seki, T.; Hagiya, M.; Shimonishi, M.; Nakamura, T.; Shimizu, S.

Gene 102, 213-219, 1991

A:Title: Organization of the human hepatocyte growth factor-encoding gene.

A:Reference number: JH0579; MUID:91340155; PMID:1831432

A:Accession: JH0579

A:Molecule type: DNA

A:Residues: 1-728 <SEK>

A:Cross-references: DBJ:D90318

A:Note: the authors translated the codon GAA for residue 662 as Gly

R:Seki, T.; Hagiya, M.; Shimonishi, M.; Nakamura, T.; Shimizu, S.

submitted to JIPID, March 1991

A:Description: Organization of the human hepatocyte growth factor-encoding gene.

A:Reference number: JH0333

A:Accession: JH0333

A:Molecule type: DNA

A:Residues: 1-481, RT', 484-728 <SE2>

R:Weidner, K.M.; Arakaki, N.; Hartmann, G.; Vandekerckhove, J.; Weingart, S.; Rieder,

Proc. Natl. Acad. Sci. U.S.A. 88, 7001-7005, 1991

A:Title: Evidence for the identity of human scatter factor and human hepatocyte growth

A:Reference number: A41140; MUID:91334393; PMID:1831266

A:Accession: A41140

A:Molecule type: mRNA

A:Residues: 1-728 <WEI>

A:Cross-references: GB:M73239; NID:g337935; PIDN:AAA64239.1; PID:g337936

R:Seki, T.; Ihara, I.; Sugimura, A.; Shimonishi, M.; Nishizawa, T.; Asami, O.; Hagiya

Biochem. Biophys. Res. Commun. 172, 321-327, 1990

A:Title: Isolation and expression of cDNA for different forms of hepatocyte growth fa

A:Reference number: A36677; MUID:91025062; PMID:2145836

A:Accession: B36677

A:Molecule type: mRNA

A:Residues: 1-728 <SE3>

A:Cross-references: GB:M60718; NID:g184031; PIDN:AAA52648.1; PID:g184032

A:Accession: A36677

A:Molecule type: mRNA

A:Residues: 1-161,167-728 <SE4>

A:Cross-references: EMBL:X16323

A:Experimental source: leukocyte

R:Miyazawa, K.; Tsubouchi, H.; Naka, D.; Takahashi, K.; Okigaki, M.; Arakaki, N.; Nak

Biochem. Biophys. Res. Commun. 163, 967-973, 1989

A:Title: Molecular cloning and sequence analysis of cDNA for human hepatocyte growth

A:Reference number: A33512; MUID:89392017; PMID:2528952

A:Accession: A33512

A>Status: not compared with conceptual translation

A:Molecule type: mRNA

A:Residues: 1-728 <MY>

A:Cross-references: GB:M29145; NID:q184041; PIDN:AAA52650.1; PID:g306846  
R:Rubin, J.S.; Chan, A.M.L.; Bottaro, D.P.; Burgess, W.H.; Taylor, W.G.; Cech, A.C.; Hlin  
Proc. Natl. Acad. Sci. U.S.A. 88, 415-419, 1991  
A:Title: A broad-spectrum human lung fibroblast-derived mitogen is a variant of hepatocyte  
A:Reference number: A39006; MUID:91110540; PMID:1824873  
A:Accession: A39006  
A:Molecule type: mRNA  
A:Residues: 1-161,167-728 <RUB>  
A:Cross-references: GB:M5379  
A:Experimental source: embryonic lung  
R:Yoshiyama, Y.; Arahaki, N.; Naka, D.; Takahashi, K.; Hirono, S.; Kondo, J.; Nakayama,  
Biochem. Biophys. Res. Commun. 175, 660-667, 1991  
A:Title: Identification of the N-terminal residue of the heavy chain of both native and  
A:Reference number: PH0114; MUID:91207365; PMID:1826837  
A:Accession: PH0114  
A:Molecule type: protein  
A:Residues: 32-43,53-58 <YOS>  
A:Experimental source: plasma  
R:Weidner, K.M.; Behrens, J.; Vandekerckhove, J.; Birchmeier, W.  
J. Cell Biol. 111, 2097-2108, 1990  
A:Title: Scatter factor: molecular characteristics and effect on the invasiveness of ep  
A:Reference number: A37796; MUID:91035621; PMID:2146276  
A:Accession: A37796  
A:Molecule type: protein  
A:Residues: 86-91,329-344; 356-363, 'XX', 366-370; 425-434; 442-447, 'X', 449-450; 543-546, 'X', 5  
R:Nakamura, T.; Nishizawa, T.; Hagiya, M.; Seki, T.; Shimonishi, M.; Sugimura, A.; Tashi  
Nature 342, 440-443, 1989  
A:Title: Molecular cloning and expression of human hepatocyte growth factor.  
A:Reference number: S06794; MUID:90066676; PMID:2531289  
A:Accession: S06794  
A:Molecule type: mRNA  
A:Residues: 1-31, 'HK', 34-77, 'N', 79-292, 'V', 294-299, 'M', 301-316, 'A', 318-335, 'K', 337-386, 'E'  
A:Cross-references: EMBL:X16323; NID:g32081; PIDN:CAA34387.1; PID:g32082  
A:Experimental source: liver  
A:Note: the authors translated the codon CAG for residue 727 as Glu  
A:Note: part of this sequence, including the amino end of both the alpha and beta chains  
R:Hartmann, G.; Naidin, L.; Weidner, K.M.; Sachs, M.; Vigna, E.; Comoglio, P.M.; Birchm  
Proc. Natl. Acad. Sci. U.S.A. 89, 11574-11578, 1992  
A:Title: A functional domain in the heavy chain of scatter factor/hepatocyte growth fact  
A:Reference number: I59214; MUID:93087571; PMID:1280830  
A:Accession: I59214  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-288, 'ET' <HAR>  
A:Cross-references: GB:I02931; NID:g184033; PIDN:AAA52649.1; PID:g184034  
R:Miyazawa, K.; Kitamura, A.; Naka, D.; Kitamura, N.  
Eur. J. Biochem. 197, 15-22, 1991  
A:Title: An alternatively processed mRNA generated from human hepatocyte growth factor  
A:Reference number: S15443; MUID:91200041; PMID:1826653  
A:Accession: S15443  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-288, 'ET' <MIY2>  
A:Cross-references: EMBL:X57574; NID:g32083; PIDN:CAA40802.1; PID:g32084  
R:Shima, N.; Nagao, M.; Ogaki, F.; Tsuda, E.; Murakami, A.; Higashio, K.  
Biochem. Biophys. Res. Commun. 180, 1151-1158, 1991  
A:Title: Tumor cytotoxic factor/hepatocyte growth factor from human fibroblasts: cloning  
A:Reference number: I52253; MUID:92062058; PMID:1835383  
A:Accession: I52253  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 161-166 <SHI>  
A:Cross-references: GB:S62561; NID:g237996; PIDN:AAB20169.1; PID:g237997  
A:Gene: GDB:HGF  
C:Genetics:  
A:Cross-references: GDB:127524; OMIM:142409  
A:Map position: 7q21.1-7q21.1  
A:Introns: 30/1; 85/2; 123/1; 161/2; 209/1; 249/2; 289/1; 347/2; 390/1; 424/2; 469/1; 48  
C:Complex: disulfide-bonded heterodimer of chains derived from the same precursor  
C:Function:  
A:Description: stimulates mitosis of hepatocytes and other cells  
A:Note: does not have proteinase activity  
C:Superfamily: hepatocyte growth factor; kringle homology; trypsin homology

C:Keywords: alternative splicing; glycoprotein; growth factor; heterodimer; kringle;  
F:1-31/Domain: signal sequence #status predicted <SIG>  
F:32-494,495-728/Product: hepatocyte growth factor #status experimental <MAT>  
F:32-494/Domain: alpha chain #status experimental <ACH>  
F:128-206/Domain: kringle homology <KR1>  
F:211-288/Domain: kringle homology <KR2>  
F:305-383/Domain: kringle homology <KR3>  
F:391-469/Domain: kringle homology <KR4>  
F:495-728/Domain: beta chain #status experimental <BCH>  
F:495-716/Domain: trypsin homology <TRY>  
F:32/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status experi  
F:294,402,566,653/Binding site: carboxydrate (Asn) (covalent) #status predicted  
F:487-604/Disulfide bonds: #status predicted

Query Match 99.9%; Score 2601; DB 1; Length 728;

Best Local Similarity 99.8%; Pred. No. 2,1e-177;

Matches 446; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ERKRNTTIEFKSAKTTLLIKIDPALKIKYKVNADOCANRCTRNGKLPFTCKAFVFK 60

Db 32 QKRKRNTTIEFKSAKTTLLIKIDPALKIKYKVNADOCANRCTRNGKLPFTCKAFVFK 91

QY 61 ARKOCLEFPENSSGVYKKEGHEFDLYENKDYIRNCIIIGKGRSYKGTVSITKSGIKCQP 120

Db 92 ARKOCLEFPENSSGVYKKEGHEFDLYENKDYIRNCIIIGKGRSYKGTVSITKSGIKCQP 151

QY 121 WSSMIPHEHSELPSSYRGKDLQENYCRNPRGEGGPMCFSTSNPEVRYEVCIDPOCSEVEC 180

Db 152 WSSMIPHEHSELPSSYRGKDLQENYCRNPRGEGGPMCFSTSNPEVRYEVCIDPOCSEVEC 211

QY 181 MTCNGESYRGLMDHTESKICORWDHOTPHRHKEFLPERYPDKGPDNNYCRNPDDQPRWC 240

Db 212 MTCNGESYRGLMDHTESKICORWDHOTPHRHKEFLPERYPDKGPDNNYCRNPDDQPRWC 271

QY 241 YLDPHTRWYCAIKTCADNTMNDTVDPLETTECIQGGEGYGTVNTIWMGIPCQRWDS 300

Db 272 YLDPHTRWYCAIKTCADNTMNDTVDPLETTECIQGGEGYGTVNTIWMGIPCQRWDS 331

QY 301 QYPHEHMTPEKFKDLRENYCRNPDCSESPWCFSTDPNIRVGYCSOIPNCDSHSGDC 360

Db 332 QYPHEHMTPEKFKDLRENYCRNPDCSESPWCFSTDPNIRVGYCSOIPNCDSHSGDC 391

QY 361 YRGNGKNYMGNLSTQRTSLGTCMMDKNNMEDLHRHIFWEPDASKLNENYCRNPDDDAHGPW 420

Db 392 YRGNGKNYMGNLSTQRTSLGTCMMDKNNMEDLHRHIFWEPDASKLNENYCRNPDDDAHGPW 451

QY 421 CYTGNPLIPWDYCPISRCGEGDTTPTIV 447

Db 452 CYTGNPLIPWDYCPISRCGEGDTTPTIV 478

## RESULT 2

A60185

hepatocyte growth factor precursor - mouse

N:Alternate names: hepalectin A; scatter factor

C:Species: Mus musculus (house mouse)

C:Date: 03-Mar-1993 #sequence revision 26-May-1994 #text change 16-Jun-2000

C:Accession: J02117; PC2064; A60185; S43416; S45521; S17173; S10966; I48758; J00231

R:Sasaki, M.; Nishio, M.; Sasaki, T.; Enami, J.

Biochem. Biophys. Res. Commun. 199, 772-779, 1994

A:Title: Identification of mouse mammary fibroblast-derived mammary growth factor as

A:Reference number: J02117; MUID:94183257; PMID:8135822

A:Accession: J02117

A:Molecule type: mRNA

A:Residues: 1-728 <SAS2>

A:Cross-references: GB:D10212; NID:g220435; PIDN:BAA01064.1; PID:g220436

A:Experimental source: fibroblast, COS-1 cell

A:Note: submitted to JIPID, May 1993

A:Accession: PC2064

A:Molecule type: protein

A:Residues: 496-504 <SAS2>

R:Rosen, E.M.; Meromsky, L.; Setter, E.; Vinter, D.W.; Goldberg, I.D.

Proc. Soc. Exp. Biol. Med. 195, 34-43, 1990



Query Match 92.2%; Score 2401; DB 1; Length 728;  
Best Local Similarity 90.4%; Pred. No. 3.5e-163;  
Matches 404; Conservative 23; Mismatches 20; Indels 0; Gaps 0;

QY 1 ERKRNTIHEFKSAKTLIKIDPALIKIKKYNTADOCANRCTRNGKGLPFTCKAFVFDK 60  
DB 33 OKKRNTIHEFKSAKTLIKIDPALIKIKKYNTADOCANRCTRNGKGLPFTCKAFVFDK 92  
QY 61 ARKQCLWFFPNSMSSGVKKFEGHEFDLYENKDYIRNCIIIGKRSYKGTVITKSGIKCQP 120  
DB 93 SRKRCYWPNSMSSGVKKFEGHEFDLYENKDYIRNCIIIGKRSYKGTVITKSGIKCQP 152  
QY 121 WSMIPHEHSLPSSYRGKDLQENYCRNPGEGGWCFTSNPEVRYEVCIDIPOCSEVEC 180  
DB 153 WSMIPHEHSLPSSYRGKDLQENYCRNPGEGGWCFTSNPEVRYEVCIDIPOCSEVEC 212  
QY 181 MTCNGESYRGLMDHTESGKICQRDWHDTPHHRKFLPERYDPKGFDDNYCRNPDGQPRWC 240  
DB 213 MTCNGESYRGLMDHTESGKICQRDWHDTPHHRKFLPERYDPKGFDDNYCRNPDGQPRWC 272  
QY 241 YLDPHTRWECYCAIKTCADNTMDTVPLETCIQOGEGYRGTVNTIWNIGPCQRWDS 300  
DB 273 YLDPHTRWECYCAIKTCADNTMDTVPLETCIQOGEGYRGTVNTIWNIGPCQRWDS 332  
QY 301 QYPHEHMTPEHFKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDMSHGDC 360  
DB 333 QYPHEHMTPEHFKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDMSHGDC 392  
QY 361 YRGNKNTYMGNSLQTRSGLTCSMDKNNMEDLHRHIFWEPDASKLNEYCRNPDGQPRWC 420  
DB 393 YRGNKNTYMGNSLQTRSGLTCSMDKNNMEDLHRHIFWEPDASKLNEYCRNPDGQPRWC 452  
QY 421 CYTGNPLIPWDYCPISRCGEGDTTPIV 447  
DB 453 CYTGNPLIPWDYCPISRCGEGDTTPIV 479

RESULT 4  
I51283  
hepatocyte growth factor precursor - clawed frog  
N:Alternate names: hepatopoietin A; scatter factor  
C:Species: Xenopus sp. (clawed frog)  
C>Date: 13-Sep-1996 #sequence\_revision 13-Sep-1996 #text\_change 18-Jun-1999  
C:Accession: I51283  
R:Nakamura, H.; Tashiro, K.; Nakamura, T.; Shiokawa, K.  
Mech. Dev. 49, 123-131, 1995  
A:Title: Molecular cloning of Xenopus HGF cDNA and its expression studies in Xenopus ear  
A:Reference number: I51283; MUID:95267690; PMID:7748783  
A:Accession: I51283  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-710 <NA>  
A:Cross-references: GB:S77422; NID:9989932; PIDN:AB34354.1; PID:g998933  
A:Note: the authors' translation for residue 458 (Thr) is inconsistent with the nucleotide  
C:Complex: disulfide-bonded heterodimer of chains derived from the same precursor  
C:Function:  
A:Description: stimulates mitosis of hepatocytes and other cells  
A:Note: does not have proteinase activity  
C:Superfamily: hepatocyte growth factor; kringe homology; trypsin homology  
F:42-477/478-709/product: hepatocyte growth factor #status predicted <NA>  
F:42-477/Domain: hepatocyte growth factor alpha chain #status predicted <ACH>  
F:115-193/Domain: kringe homology <KR1>  
F:198-275/Domain: kringe homology <KR2>  
F:289-367/Domain: kringe homology <KR3>  
F:375-453/Domain: kringe homology <KR4>  
F:478-709/Domain: hepatocyte growth factor beta chain #status predicted <BCH>  
F:478-709/Domain: trypsin homology <TRY>  
F:52,128,281,322,379,550,637,666/Binding site: carbohydrate (Asn) (covalent) #status pre  
F:470-588/Disulfide bonds: #status predicted

Query Match 71.6%; Score 1864.5; DB 1; Length 710;

Best Local Similarity 69.3%; Pred. No. 4.8e-125;  
Matches 305; Conservative 64; Mismatches 68; Indels 3; Gaps 1;

QY 3 KRNTIHEFKSAKTLIKIDPALIKIKKYNTADOCANRCTRNGKGLPFTCKAFVFDKAR 62  
DB 21 KRNNAFDDYKTAETTLRLNKALEVKTKMFNTTENCARCSRNGKGLPFTCKAFADKNI 80  
QY 63 KQCLWFFPNSMSSGVKKFEGHEFDLYENKDYIRNCIIIGKRSYKGTVITKSGIKCQPS 122  
DB 81 KRCHWFSPTMSAGIKDYISFDLYEKDYIRDCIHGKGSNYRGTNRNVRKGLACQPN 140  
QY 123 SMIPHEHSLPSSYRGKDLQENYCRNPGEGGWCFTSNPEVRYEVCIDIPOCSEVECMT 182  
DB 141 SMIPHEHSLPSSYRGKDLQENYCRNPGEGGWCFTSNPEVRYEVCIDIPOCSEVECMT 200  
QY 183 CNGESYRGLMDHTESGKICQRDWHDTPHHRKFLPERYDPKGFDDNYCRNPDGQPRWCYT 242  
DB 201 CNGESYRGLMDHTESGKICQRDWHDTPHHRKFLPERYDPKGFDDNYCRNPDGQPRWCYT 260  
QY 243 LDPHTRWECYCAIKTCADNTMDTVPLETCIQOGEGYRGTVNTIWNIGPCQRWDSOY 302  
DB 261 LDPHTRWECYCAIKTCADNTMDTVPLETCIQOGEGYRGTVNTIWNIGPCQRWDSOY 317  
QY 303 PHEHMTPEHFKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDMSHGDCYR 362  
DB 318 PHEHMTPEHFKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDMSHGDCYR 377  
QY 363 GNGKNTYMGNSLQTRSGLTCSMDKNNMEDLHRHIFWEPDASKLNEYCRNPDGQPRWCYT 422  
DB 378 GNGKNTYMGNSLQTRSGLTCSMDKNNMEDLHRHIFWEPDASKLNEYCRNPDGQPRWCYT 437  
QY 423 TGNPLIPWDYCPISRCGEGDT 442  
DB 438 TGNPLIPWDYCPISRCGEGDT 457

RESULT 5  
I51285  
hepatocyte growth factor/scatter factor - chicken (fragment)  
C:Species: Gallus gallus (chicken)  
C>Date: 13-Sep-1996 #sequence\_revision 13-Sep-1996 #text\_change 16-Jul-1999  
C:Accession: I51285  
R:Streit, A.; Stern, C.D.; Thery, C.; Ireland, G.W.; Aparicio, S.; Sharpe, M.J.; Gher  
Development 121, 813-824, 1995  
A:Title: A role for HGF/SF in neural induction and its expression in Hensen's node du  
A:Reference number: I51285; MUID:95237013; PMID:7720585  
A:Accession: I51285  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-411 <STR>  
A:Cross-references: GB:S77480; NID:g998675; PID:g998676  
C:Superfamily: hepatocyte growth factor; kringe homology; trypsin homology  
F:124-197/Domain: kringe homology <KR>  
F:202-279/Domain: kringe homology <KR2>  
F:296-374/Domain: kringe homology <KR3>

Query Match 66.7%; Score 1737.5; DB 2; Length 411;  
Best Local Similarity 74.9%; Pred. No. 2.9e-116;  
Matches 289; Conservative 47; Mismatches 45; Indels 5; Gaps 1;

QY 3 KRNTIHEFKSAKTLIKIDPALIKIKKYNTADOCANRCTRNGKGLPFTCKAFVFDKAR 62  
DB 30 KRNPLHDYKKTGELMLIKYNTLEVKTLNTEQCAKCSRNGKGLSFTCKAFAYDRVT 89  
QY 63 KQCLWFFPNSMSSGVKKFEGHEFDLYENKDYIRNCIIIGKRSYKGTVITKSGIKCQPS 122  
DB 90 KRCHWLSNLTNGVRKQDHAFLFEKDYVRNCIIIGKAEYKGTISITKSGIQCAWN 149  
QY 123 SMIPHEHSLPSSYRGKDLQENYCRNPGEGGWCFTSNPEVRYEVCIDIPOCSEVECMT 182  
DB 150 SMIPHEHSLPSSYRGKDLQENYCRNPGEGGWCFTSNPEVRYEVCIDIPOCSEVECMT 204  
QY 183 CNGESYRGLMDHTESGKICQRDWHDTPHHRKFLPERYDPKGFDDNYCRNPDGQPRWCYT 242

Qy	33	VNTADQCANRCTRNKGLPFTCKAFVDFDKARQCLWLPFPNSMSGVKKFEHGFDLYENKD	92
Db	50	VADAEACAGRC---GPLMDCRAFYHNVSSHGCOLLPWTQHSPTHLRLRRSGRCDLFQKDK	105
Qy	93	YIRNCIIKGKRSYKGTVSITKSGIKCPQWSSMIPHEHSFLPSSYRGKDLQENYCRNPRGE	152
Db	106	YVRTCINNGVGYRGVTWATTVGGLPCQANSHKFPNDHKYPTLRNG--LEENFCRNPDDG	163
Qy	153	EGGPWCFTSNPEVRYEYCDIPQCSEVECMTCNGESYRGLMDHTESGKICQRWDHQTPHRH	212
Db	164	PGGPWCYTTDPVAFRFQSCGIGKSCREAAVCWNGEYRGAVDRTESGRCQWDLQHPQH	223
Qy	213	KFLPERYPDKGFDNNYCRNPDGGRPMWCYTLDPHTRHEYCAIKTCADNTMNDTDVPLETT	272
Db	224	PFEFGKFLDGLDNNYCRNPDGGRPMWCYTTDQIREFCDLPRGSEAOQROEA--TTV	281
Qy	273	ECIOGOCGEYRGVTWTWNGIPCORWDSQYPHEHDMTPENFKCKDLRENYCRNPDGSESP	332
Db	282	SCFRGKGEYRGVTANTTAVGPCORWDAQIPIQHRTFPEKYACKOLRENYCRNPDGSEAP	341
Qy	333	WCFTTDPNIRVGYCSQIPNC--DMSGODQYRGNGKNTYMGNLSTFRSLGTCSMMDKNMEDL	391
Db	342	WCFTLRPMGMAAFQYQIRRCRTDDVPQDCYHGAGEYRGVTSKTRKGVCQQRWSAETPHK	401
Qy	392	HRHLEWEPDASKLNNYCRNPDGDDAHGPWCYTGNNLPWDCYPISRCEGDTTIV	447
Db	402	POFTTSEPHQAQLEENFCRNPDDGSHGPWCYTMDPRTFDFYCALRRACDDOPPSIL	457
RESULT 7			
JC5061			
macrophage-stimulating protein 1 precursor - rat			
C:Species: Rattus norvegicus (Norway rat)			
C:Date: 31-Jan-1997 #sequence_revision 31-Jan-1997 #text_change 16-Jun-2000			
C:Accession: J05061			
R:Oshiro, K.; Iwama, A.; Matsuno, K.; Ezaki, T.; Sakamoto, O.; Hamaguchi, I.			
Biochem. Biophys. Res. Commun. 227, 273-280, 1996			
A:Title: Molecular cloning of rat macrophage-stimulating protein and its i			
A:Reference number: JC5061; MUID:97011126; PMID:8858136			
A:Accession: JC5061			
A:Molecule type: mRNA			
A:Residues: 1-716 <OH>			
A:Cross-references: EMBL:X95096; NID:g1669718; PIDN:CAA64473.1; PID:g16699			
C:Complex: disulfide-bonded heterodimer of chains derived from the same pr			
C:Superfamily: hepatocyte growth factor; kringle homology; trypsin homology			
C:Keywords: duplication; glycoprotein; growth factor; kringle			
F:1-31/Domain: signal sequence #status predicted <SIG>			
F:32-488/489-716/Product: macrophage-stimulating protein 1 #status predict			
F:32-488/Domain: macrophage-stimulating protein 1 alpha chain #status prec			
F:110-186/Domain: kringle homology <KR1>			
F:191-268/Domain: kringle homology <KR12>			
F:252-370/Domain: kringle homology <KR13>			
F:379-457/Domain: kringle homology <KR14>			
F:489-716/Domain: macrophage-stimulating protein 1 beta chain #status prec			
F:489-709/Domain: trypsin homology <TRY>			
F:72,305,620/Binding site: carbohydrate (Asn) (covalent) #status predicted			
Query Match 43.9%; Score 1144; DB 1; Length 716;			
Best Local Similarity 46.1%; Pred. No. 8.3e-74;			
Matches 196; Conservative 70; Mismatches 141; Indels 18; Gaps			
Qy	33	VNTADQCANRCTRNKGLPFTCKAFVDFDKARQCLWLPFPNSMSGVKKFEHGFDLYENKD	92
Db	50	VADAEACAGRC---GPLMDCRAFYHNVSSHGCOLLPWTQHSPTHLRLRRSGRCDLFQKDK	105
Qy	93	YIRNCIIKGKRSYKGTVSITKSGIKCPQWSSMIPHEHSFLPSSYRGKDLQENYCRNPRGE	152
Db	106	YVRTCINNGVGYRGVTWATTVGGLPCQANSHKFPNDHKYPTLRNG--LEENFCRNPDDG	163
Qy	153	EGGPWCFTSNPEVRYEYCDIPQCSEVECMTCNGESYRGLMDHTESGKICQRWDHQTPHRH	212
Db	164	PGGPWCYTTNRSVRFQSCGIGKSCREAVYCWNGEDYRGVDPVETSGRCQWDLQHPHSH	223
Qy	213	KFLPERYPDKGFDNNYCRNPDGGRPMWCYTLDPHTRHEYCAIKTCADNTMNDTDVPLETT	272
Db	224	PFEFGKFLDGLDNNYCRNPDGGRPMWCYTTDQIREFCDLPRGSEAOQROEA--TTV	281
Qy	273	ECIOGOCGEYRGVTWTWNGIPCORWDSQYPHEHDMTPENFKCKDLRENYCRNPDGSESP	332
Db	282	SCFRGKGEYRGVTANTTAVGPCORWDAQIPIQHRTFPEKYACKOLRENYCRNPDGSEAP	341
Qy	333	WCFTTDPNIRVGYCSQIPNC--DMSGODQYRGNGKNTYMGNLSTFRSLGTCSMMDKNMEDL	391
Db	342	WCFTLRPMGMAAFQYQIRRCRTDDVPQDCYHGAGEYRGVTSKTRKGVCQQRWSAETPHK	401
Qy	392	HRHLEWEPDASKLNNYCRNPDGDDAHGPWCYTGNNLPWDCYPISRCEGDTTIV	447
Db	402	POFTTSEPHQAQLEENFCRNPDDGSHGPWCYTMDPRTFDFYCALRRACDDOPPSIL	457

213	QY	KELPERYPDKGEDDNYCRNPDGQPPRWCVTLDPHTRWEYCAIKTKCADN-----TMNDT	265
224	Db	PFQEFKDKLDKNDYCRNPDGSERPWCYTTPDNEVEREFCDLPSGPNLPPTVKGSKSR	283
266	QY	DVPLETTECLOGEGEYRGFTVNTWNGIPQQRWDSQYPHEHDWTPENFKCKDLRENYCRN	325
284	Db	RNKGKALNCFRGKGEYRGTTNTTSAGVFCQQRWDAQSPHQHREVPBKYACKDLRENFEN	343
326	QY	PDGSESPWCFTTDPNIRVGYCSQIPNC-DMSHGQDCYRGNGKNYMNLSOTRSLGTCSSW	384
344	Db	PDGSEAPWCFTSRPLGRMAFCHQIPRCTEELVPFGCYHGSGEQYRGVSVKTRAGVOCQHW	403
385	QY	DKNWEDLHRHIFWEPASK--LNENYCRNPDDDAHGPWCYTGNPLLPWPYCYPTSRCEGT	442
404	Db	--SSEITHPKPQFTPTSPAPQNGLEANFCRNPDGDSHGFPWCYTLPDPDLFDYCALQRCDQDQ	461
443	QY	TPTIV	447
462	Db	PPSIL	466

## RESULT 9

PLC8  
plasmin (EC 3.4.21.7) precursor - pig (fragment)  
N:Alternate names: plasminogen  
N:Contains: miniplasminogen  
C:Species: Sus scrofa domestica (domestic pig)  
C>Date: 07-Sep-1990 #sequence\_revision 01-Nov-1996 #text\_change 18-Jul-1997  
C:Accession: S03733; S03737; A25834  
R:Schaller, J.; Marti, T.; Roessellet, S.J.; Kaempfer, U.; Rickli, E.E.  
Fibrinolysis 1, 91-102, 1987  
A:Title: Amino acid sequence of the heavy chain of porcine plasmin. Comparison of the  
A:Reference number: S03733  
A:Accession: S03733  
A:Molecule type: protein  
A:Residues: 1-560 <SCH>  
R:Brundish, R.A.; Lerch, P.G.; Schaller, J.; Rickli, E.E.; Lergier, W.; Manneberg, E.  
Eur. J. Biochem. 114, 465-470, 1981  
A:Title: Comparison of the primary structure of the N-terminal CNBr fragments of human  
A:Reference number: S03735; MUID:81212097; PMID:7238497

A:Molecule type: protein  
A:Residues: 1-57 <BRU>  
R:Marti, T.; Schaller, J.; Rickli, E.E.  
Eur. J. Biochem. 149, 279-285, 1985  
A:Title: Determination of the complete amino-acid sequence of porcine miniplasminogen  
A:Reference number: A25834; MUID:85203907; PMID:3846533  
A:Accession: A25834  
A:Molecule type: protein  
A:Residues: 450-790 <MAR>  
C:Function:  
A:Description: dissolves the fibrin of blood clots; acts as a proteolytic factor in a  
ns the walls of the graafian follicle; also activates the urokinase-type plasminogen  
A:Pathway: fibrinolysis  
C:Superfamily: plasmin; kringle homology; plasminogen-related protein precursor homol  
C:Keywords: fibrinolysis; glycoprotein; hydrolase; kidney; kringle; plasma; serine pr  
F:1-790/Product: plasminogen #status predicted <PRO>  
F:1-77/Domain: plasminogen-related protein precursor homology (fragment) <PLPH>  
F:1-77/Domain: activation peptide #status predicted <APT>  
F:78-560/Product: plasmin chain A #status predicted <ACH>  
F:84-162/Domain: kringle homology <KR1>  
F:166-243/Domain: kringle homology <KR2>  
F:256-333/Domain: kringle homology <KR3>  
F:338-435/Domain: kringle homology <KR4>  
F:450-790/Product: miniplasminogen #status experimental <MIN>  
F:461-540/Domain: kringle homology <KR5>  
F:561-790/Product: plasmin chain B #status experimental <BCH>  
F:561-783/Domain: trypsin homology <TRY>  
F:30-54,34-42,84-162,105-145,133-157,166-243,169-297,187-226,215-238,256-333,277-316,  
bonds: #status predicted  
F:602,645,740/Active site: His, Asp, Ser #status predicted

Query Match 40.98; Score 1066; DB 1; Length 790;

## Query Match

Best Local Similarity 44.6%, Pred. No. 3.2e-68;  
Matches 195; Conservative 65; Mismatches 137; Indels 40; Gaps 13;

Qy 25 ALIKITKVVNTADOCANCRTRNKLGLPFTCKAFVDFDKARKQCLWPFNSMSGVKKKEFGHE 84  
Db 16 SLRSKQVAARSVEECAACEAETN--FCRAFOYHSKQOQCVVMAENSKTSPARM--RD 71

Qy 85 FDLYENKDYIRNCIITGKRSYKGTVSTIKSGIKQOPWSSMIPHEHSFLPSSYRGKDLQEN 144  
Db 72 VVLFKRIYLSCKTGTGNGKNTGTSKTSKGVICOKWSVSSPHIPKYSPEKFLAGLEEN 131

Qy 145 YCRNPRGEGGPWCFTSNPEVRYEVCIDIPOQSEVECTNCESYRGLMDHTESGKICQW 204  
Db 132 YCRNPNDKEGKWCYTTDPTREDFYCDIPEC--EDECMHCSGEHYEGKISKTMGSLGECOSW 190

Qy 205 DHQTHRHKLPERYPDKGFDNDNYCRNPDPGPRWCYITLDPHTRWECYCAIKTCADNTWND 264  
Db 191 GSOSPHAGLYLSPKPNKLNMYCRNPDGEPKWCFTTDPNKRWFCDIPRC--TTPPP 248

Qy 265 TDVPLETTECQOGGEGYRGTVNTWINGIPQORWDSQYPHEHDMTPENFKCKDLRENYCR 324  
Db 249 TSGP--TYQCLKGRGENYRGTVSVTASGHTCQWQAQSPHKNRTPENFPCKNLEENYCR 306

Qy 325 NPDGSESPWCFTTDNIRVGYCSQIPNCDMS-----H-----GQDCYRGNGKN 367  
Db 307 NPDGETAPWCYTTDSEVRWDYC-KIPSCGSSSTSTEHLDAVPVPBQTPVAQDCYRGNGES 365

Qy 368 YGNLSQTSRSLGTCMWDKNMEDLHRHFWF-----PDASKLNENYCRNPDDAHGPWC 421  
Db 366 YRGTSSTITGRKCSQWVSMT--HRH---EKTFCGNFFNAG--LTWNYCRNPDD--KSPWC 418

Qy 422 YTGPNLPDWCPISRC 438  
Db 419 YTTDPVRWEYCNLKKC 435

RESULT 10  
PLMS  
plasmin (EC 3.4.21.7) precursor - mouse  
N:Contains: angiotatin; plasminogen  
C:Species: Mus musculus (house mouse)  
C:Date: 20-Sep-1991 #sequence\_revision 01-Nov-1996 #text\_change 18-Jun-1999  
C:Accession: A38514; S48202; S48203  
R:Deegen, S.J.F.; Bell, S.M.; Schaefer, L.A.; Elliott, R.W.  
Genomics 8, 49-61, 1990  
A:Title: Characterization of the cDNA coding for mouse plasminogen and localization of the  
A:Reference number: A38514; MUID:91184812; PMID:2081600  
A:Accession: A38514  
A:Molecule type: mRNA  
A:Residues: 1-812 <DEG>  
A:Cross-references: GB:J04766; NID:g200402; PIDN:AAA50168.1; PID:g200403  
R:Lijnen, H.R.; van Hoef, B.; Beelen, V.; Collen, D.  
Eur. J. Biochem. 224, 863-871, 1994  
A:Title: Characterization of the murine plasma fibrinolytic system.  
A:Reference number: S48202; MUID:95010076; PMID:7523120  
A:Accession: S48202  
A:Molecule type: protein  
A:Residues: 20-25 <LIJ>  
A:Accession: S48203  
A:Molecule type: protein  
A:Residues: 22-27 <LIJ>  
C:Comment: Plasminogen is synthesized by the kidney and is present in plasma and many of  
C:Comment: plasminogen is converted into plasmin by plasminogen activators, both plasmin  
immediately after dissociation from the clot. In the presence of the inhibitor, the activa  
e inhibitor, the activation involves also removal of the activation peptide.  
C:Comment: Stromelysin 1 (see PIR:KCMSL) acts on plasminogen to produce angiotatin. TC  
eul in treating solid tumors.  
C:Function:  
A:Description: dissolves the fibrin of blood clots; acts as a proteolytic factor in a va  
ns the walls of the graafian follicle; also activates the urokinase-type plasminogen act  
A:Pathway: fibrinolysis  
C:Superfamily: plasmin; kringle homology; plasminogen-related protein precursor homology  
C:Keywords: angiotatin inhibitor; blood; duplication; fibrinolysis; glycoprotein; hyd

F:1-96/Domain: plasminogen-related protein precursor homology <PLPH>  
F:1-19/Domain: signal sequence #status predicted <SIG>  
F:20-812/Product: plasminogen #status predicted <PRO>  
F:20-96/Domain: activation peptide #status predicted <APT>  
F:79-466/Product: angiotatin #status predicted <AST>  
F:97-581,582-812/Product: plasmin #status predicted <MAT>  
F:97-581/Domain: chain A #status predicted <ACH>  
F:103-181/Domain: kringle homology <KR1>  
F:185-262/Domain: kringle homology <KR2>  
F:275-352/Domain: kringle homology <KR3>  
F:377-454/Domain: kringle homology <KR4>  
F:481-560/Domain: kringle homology <KR5>  
F:582-812/Domain: chain B #status predicted <BCH>  
F:582-805/Domain: trypsin homology <TRY>  
F:49-73,53-61,103-181,124-164,152-176,185-262,188-316,206-245,234-257,275-352,296-335  
bonds: #status predicted  
F:78-79/Cleavage site: Glu-Asn (stromelysin 1) #status predicted  
F:136,308/Binding site: carbohydrate (Asn) (covalent) #status predicted  
F:466-467/Cleavage site: Thr-Val (stromelysin 1) #status predicted  
F:581-582/Cleavage site: Arg-Val (plasminogen activator) #status experimental  
F:624,667,762/Active site: His, Asp, Ser #status predicted

Query Match 40.7%; Score 1059; DB 1; Length 812;  
Best Local Similarity 43.3%; Pred. No. 1e-67;  
Matches 191; Conservative 67; Mismatches 145; Indels 38; Gaps 12;

Qy 30 TKKVNTA---DQCANCRTRNKLGLPFTCKAFVDFDKARKQCLWPFNSMSGVKKKEFGHEFD 86  
Db 37 TKQLAAGVSDCLAKCEGE--TDFVCRSFQYHSKEQOCQVMAENSKTSSIIRM--RDVI 92

Qy 87 LYENKDYIRNCIITGKRSYKGTVSTIKSGIKQOPWSSMIPHEHSFLPSSYRGKDLQENY 146  
Db 93 LFEKRVYLSCKTGTGNGYRGTVSRTKSGVACQKNGATFPHVNYSPSTHPNEGLEENY 152

Qy 147 RNPRGEGGPWCFTSNPEVRYEVCIDIPOQSEVECTNCESYRGLMDHTESGKICQW 206  
Db 153 RNPDNDKEGKWCYTTDPTREDFYCDIPEC--ECYSGEKYEGKISKTMGLDQOAWDS 211

Qy 207 QTPHRHKLPERYPDKGFDNDNYCRNPDPGPRWCYITLDPHTRWECYCAIKTCADNTW 266  
Db 212 QSPHAGYIPAKFPKSNKLNMYCRNPDGEPKWCFTTDPNKRWFCDIPRC--TTPPP 267

Qy 267 VPLETTECQOGGEGYRGTVNTWINGIPQORWDSQYPHEHDMTPENFKCKDLRENYCR 326  
Db 268 PPSPTYQCLKGRGENYRGTVSVTASGHTCQWQAQSPHKNRTPENFPCKNLEENYCR 327

Qy 327 DGSSESPWCFTTDNIRVGYCSQIPNCDMSHG-----ODCYRGNGKNY 369  
Db 328 DGETAPWCYTTDSEVRWDYC-KIPSCGSSSTSTEHLDAVPVPBQTPVAQDCYRGNG 386

Qy 370 GNLSQTSRSLGTCMWDKNMEDLHRHFWF-----PDASKLNENYCRNPDDAHGPWCY 426  
Db 387 GTSSTITGRKCSQWVSMT--HRHKTPEFPDAG--LEMNYCRNPDDG--KGPWCYTTDP 442

Qy 427 LIPWDYCPISRCGDTTPTIV 447  
Db 443 SVRWEYCNLKR--SETGGSV 462

RESULT 11  
I46260  
plasmin (EC 3.4.21.7) precursor - western European hedgehog  
C:Species: Erinaceus europaeus (western European hedgehog)  
C:Date: 21-Feb-1997 #sequence\_revision 21-Feb-1997 #text\_change 16-Jul-1999  
C:Accession: I46260  
R:Lawn, R.M.; Boommark, N.W.; Schwartz, K.; Lindahl, G.E.; Wade, D.P.; Byrne, C.D.; F  
J. Biol. Chem. 270, 24004-24009, 1995  
A:Title: The recurring evolution of Lp(a): Insights from cloning of hedgehog apolipop  
A:Reference number: I46259; MUID:96025778; PMID:7592597  
A:Accession: I46260  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-810 <LAW>



A:Cross-references: EMBL:U33171; NID:g1046360; PID:g1046361  
C:Superfamily: plasmin; kringle homology; plasminogen-related protein precursor homology  
C:Keywords: hydrolase; serine proteinase  
F:1-96/Domain: plasminogen-related protein precursor homology <PLPH>  
F:103-181/Domain: kringle homology <KR1>  
F:185-262/Domain: kringle homology <KR2>  
F:275-352/Domain: kringle homology <KR3>  
F:379-456/Domain: kringle homology <KR4>  
F:482-561/Domain: kringle homology <KR5>  
F:582-803/Domain: trypsin homology <TRY>

Query Match 40.5%; Score 1053.5; DB 2; Length 810;  
Best Local Similarity 44.3%; Pred. No. 2.6e-67;  
Matches 194; Conservative 63; Mismatches 138; Indels 43; Gaps 13;

QY 30 TKK--VNTADOCANRCTRNGKLPFTCKAFVDFKARKQCLWFFPNSMSSGVYKKEGHEFD 86  
DB 37 TKQLSVGSTEECAVKC--EKETSFICRSFYHSKEQOCVMAENRKSIVFRM--RDV 92  
QY 87 LYENKDYIRNCIIIGKRSYKGTVSITKSGIKQCPWSSMIPHEHSFLPSSYRGKDLQENYC 146  
DB 93 LFEKKMYLSECKVNGKYYRGVTSKTLGTCQKWSAETPHKPRESPDENPSEGLDQNYC 152  
QY 147 RNPREGGPGWCFSTNPEVRYEVCIDIPQCSVEVCMTNGESYRGLMDHTESGKICQWDH 206  
DB 153 RNPNDPKPGWCYTMDEPRVYEVCEIIQC--EDECHMCSGQNTVVGKISRMTSGLEQPWD 211  
QY 207 QTPHRHKLFLPERYPDKGFDNDCRNPDGPRPWCYTLDPHTRWEYCAIKTCADNTMDTD 266  
DB 212 QIPPHGFIPTSPKPNKMLKKNYCRNPDGPRPWCFTMDNKRWEYCDIPRCT-----TP 265  
QY 267 VPLE--TTECIOGOGYRGTVNTIWNIGPCQWDSQYRPHEDMTPENFKCKDLRENYCR 324  
DB 266 PPSGGTYCQLMGNGEYHGVNVAVTYVSLGTCQWGEQSPHRDRTPENYPCKNLDENYCR 325  
QY 325 NPDGSESPWCFTTDPNIRVGYCSQIPNC-----DMSHG-----QDCYRGNG 365  
DB 326 NPDGEPAPWCFTTNSVRWEFC--KIPDCVSSASSETSDESDAPVTPPEQPPVVOECYQNG 384  
QY 366 KNYGMNLSQTRSLGTCMSMDKMDLHRHIFWE--PDASKLNENYCRNPDADDHGPWCY 422  
DB 385 QTYRGTSSTITGKKOPWTSMRP--HRHSKTPENYPDAD-LTMNCRNPDGD-KGPWCY 440  
QY 423 TGNPLIPWDYCPISRCGE 440  
DB 441 TTDPSVRWEYCNLKKCSG 458

RESULT 12  
B30848  
plasmin (EC 3.4.21.7) precursor - rhesus macaque  
C:Species: Macaca mulatta (rhesus macaque)  
C:Date: 31-Mar-1989 #sequence\_revision 31-Mar-1989 #text\_change 22-Jun-1999  
C:Accession: B32869; B30848  
R:Tromlinson, J.E.; McLean, J.W.; Lawn, R.M.  
J. Biol. Chem. 264, 5957-5965, 1989  
A:Title: Rhesus monkey apolipoprotein(a). Sequence, evolution, and sites of synthesis.  
A:Reference number: A32869; MUID:89174660; PMID:2925643  
A:Accession: B32869  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-810 <TOM>  
A:Cross-references: GB:J04697; NID:g342272; PIDN:AAA36901.1; PID:g342273  
C:Superfamily: plasmin; kringle homology; plasminogen-related protein precursor homology  
C:Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; serine proteinase  
F:1-96/Domain: plasminogen-related protein precursor homology <PLPH>  
F:103-181/Domain: signal sequence #status predicted <SIG>  
F:185-262/Domain: kringle homology <KR1>  
F:275-352/Domain: kringle homology <KR2>  
F:379-456/Domain: kringle homology <KR3>  
F:481-560/Domain: kringle homology <KR4>  
F:581-803/Domain: trypsin homology <TRY>

A:Cross-references: EMBL:U33171; NID:g1046360; PID:g1046361  
C:Superfamily: plasmin; kringle homology; plasminogen-related protein precursor homology  
C:Keywords: hydrolase; serine proteinase  
F:1-96/Domain: plasminogen-related protein precursor homology <PLPH>  
F:103-181/Domain: kringle homology <KR1>  
F:185-262/Domain: kringle homology <KR2>  
F:275-352/Domain: kringle homology <KR3>  
F:379-456/Domain: kringle homology <KR4>  
F:482-561/Domain: kringle homology <KR5>  
F:582-803/Domain: trypsin homology <TRY>

F:49-73,53-61,103-181,124-164,152-176,185-262,188-316,206-245,234-257,275-352,296-335  
bonds: #status predicted  
F:622,665,760/Active site: His, Asp, Ser #status predicted

Query Match 40.1%; Score 1044.5; DB 2; Length 810;  
Best Local Similarity 42.6%; Pred. No. 1.1e-66;  
Matches 188; Conservative 74; Mismatches 128; Indels 51; Gaps 12;

QY 30 TKK--VNTADOCANRCTRNGKLPFTCKAFVDFKARKQCLWFFPNSMSSGVYKKEGHEFD 86  
DB 37 TKQLGAGSIECAKCEEE--EFTCRSFQYHSKEQOCVMAENRKSIVFRM--RDV 92  
QY 87 LYENKDYIRNCIIIGKRSYKGTVSITKSGIKQCPWSSMIPHEHSFLPSSYRGKDLQENYC 146  
DB 93 LFEKKVYLSECKTNGKKNYRGTSKTRGITCQKWSSTSPHRPTSPATPSEGLEENYC 152  
QY 147 RNPREGGPGWCFSTNPEVRYEVCIDIPQCSVEVCMTNGESYRGLMDHTESGKICQWDH 206  
DB 153 RNPNDGPGWCYTDPPEEREDYCDIPEC--EDECHMCSGENDGKISKTMSGLEQAWDS 211  
QY 207 QTPHRHKLFLPERYPDKGFDNDCRNPDGPRPWCYTLDPHTRWEYCAIKTCADNTMDTD 266  
DB 212 QSPHAHGYIPSEKFNKLNKKNYCRNPDGPRPWCFTTDPNKRWEYCDIPRCT-----TP 265  
QY 267 VPLE--TTECIOGOGYRGTVNTIWNIGPCQWDSQYRPHEDMTPENFKCKDLRENYCR 324  
DB 266 PPSGGTYCQLMGNGEYHGVNVAVTYVSLGTCQWGEQSPHRDRTPENYPCKNLDENYCR 325  
QY 325 NPDGSESPWCFTTDPNIRVGYCSQIPNC-----DMSHG-----QDCYRGNGKN 367  
DB 326 NPDGEPAPWCFTTNSVRWEYCNLKKCSG 458  
QY 368 KNYGMNLSQTRSLGTCMSMDKMDLHRHIFWE--PDASKLNENYCRNPDADDHGP 419  
DB 385 QTYRGTSSTITGKKOPWTSMRP--HRHSKTPENYPDAD-LTMNCRNPDAD-KGP 435  
QY 420 WCYTGNPLIPWDYCPISRCGE 440  
DB 436 WCFITDPSVRWEYCNLKKCSG 456

## RESULT 13

## PLHU

plasmin (EC 3.4.21.7) precursor [validated] - human  
N:Alternate names: plasminogen precursor [misnomer]  
N:Contains: angiotatin; microplasmin; plasminogen  
C:Species: Homo sapiens (man)

C:Date: 24-Apr-1984 #sequence\_revision 02-Dec-1994 #text\_change 15-Sep-2000  
C:Accession: A35229; I52242; A26646; I62738; I84609; S03735; A00929; A04627; A04625;  
R:Petersen, T.E.; Martzen, M.R.; Ichinose, A.; Davie, E.W.  
J. Biol. Chem. 265, 6104-6111, 1990  
A:Title: Characterization of the gene for human plasminogen, a key proenzyme in the f  
A:Reference number: A35229; MUID:90202879; PMID:2318848

A:Accession: A35229

A:Molecule type: DNA

A:Residues: 1-810 <PE>

A:Cross-references: GB:J05286; GB:M34276; NID:g190064; PIDN:AAA60113.1; PID:g387026  
A:Experimental source: leukocyte; lung fibroblast  
R:Maigaretta, N.; Bruno, L.; Pontoglio, M.; Candiani, G.; Meroni, G.; Ottolenghi, S.;  
Biochem. Biophys. Res. Commun. 173, 1013-1018, 1990  
A:Title: Definition of the transcription initiation site of human plasminogen gene in  
A:Reference number: I52242; MUID:91097523; PMID:2268308

A:Accession: I52242

A:Status: translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-16 <MALI>

A:Cross-references: GB:M62890; NID:g190092; PIDN:AAA36454.1; PID:g553613

R:Forssgren, M.; Raden, B.; Israelsson, M.; Larsson, K.; Heden, L.O.

FEBS Lett. 213, 254-260, 1987

A:Title: Molecular cloning and characterization of a full-length cDNA clone for human  
A:Reference number: A26646; MUID:87162490; PMID:3030813

A:Accession: A26646

A:Molecule type: mRNA



- A:Residues: 1-471,'D',473-810 <FOR>  
A:Cross-references: GB:X05199; NID:g35530; PIDN:CAA28831.1; PID:g35531  
A:Experimental source: liver  
R:Malinowski, D.P.; Sadler, J.E.; Davie, E.W.  
Biochemistry 23, 4243-4250, 1984  
A:Title: Characterization of a complementary deoxyribonucleic acid coding for human and  
A:Reference number: 145961; MUID:85023311; PMID:6148961  
A:Accession: 162738  
A>Status: translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 292-471,'D',473-810 <MAL2>  
A:Cross-references: GB:K02922; NID:g190112; PIDN:AAA60124.1; PID:g387031  
A:Accession: 184609  
A>Status: translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 367-419 <MAL3>  
A:Cross-references: GB:K02921; NID:g190110; PIDN:AAA60123.1; PID:g190111  
R:Brünsholz, R.A.; Lerch, P.G.; Schaller, J.; Rickli, E.E.; Lergier, W.; Manneberg, M.;  
Eur. J. Biochem. 114, 463-470, 1981  
A:Title: Comparison of the primary structure of the N-terminal CNBR fragments of human,  
A:Reference number: S03735; MUID:81212097; PMID:7238497  
A:Accession: S03735  
A:Molecule type: protein  
A:Residues: 20-71,'E',73-85,87-106,'D',108-360,'E',362-810 <SOT>  
R:Wiman, B.  
R:Sottrop-Jensen, L.; Petersen, T.E.; Magnusson, S.  
submitted to the Atlas, July 1977  
A:Reference number: A00929  
A:Accession: A00929  
A:Molecule type: protein  
A:Residues: 20-71,'E',73-85,87-106,'D',108-360,'E',362-810 <SOT>  
R:Wiman, B.  
Eur. J. Biochem. 76, 129-137, 1977  
A:Title: Primary structure of the B-chain of human plasmin.  
A:Reference number: A04627; MUID:77225245; PMID:142009  
A:Accession: A04627  
A:Molecule type: protein  
A:Residues: 581-810 <MIL>  
R:Wiman, B.; Wallen, P.  
Eur. J. Biochem. 50, 489-494, 1975  
A:Title: Structural relationship between "glutamic acid" and "lysine" forms of human plasminogen  
A:Reference number: A04625; MUID:75093329; PMID:122932  
A:Accession: A04625  
A:Molecule type: protein  
A:Residues: 20-50,'Q',51-71,'E',73-85,87-100 <W12>  
R:Wiman, B.; Wallen, P.  
Eur. J. Biochem. 58, 539-547, 1975  
A:Title: Amino-acid sequence of the cyanogen-bromide fragment from human plasminogen the  
A:Reference number: A04626; MUID:76043692; PMID:126863  
A:Accession: A04626  
A:Molecule type: protein  
A:Residues: 483-507,'E',509-604 <W13>  
R:Robbins, K.C.; Bernabe, P.; Arzadon, L.; Summaria, L.  
J. Biol. Chem. 248, 1631-1633, 1973  
A:Title: The primary structure of human plasminogen. II. The histidine loop of human plasminogen  
A:Reference number: A92125; MUID:73149248; PMID:4694729  
A:Contents: annotation: active site  
R:Groskopf, W.R.; Summaria, L.; Robbins, K.C.  
J. Biol. Chem. 244, 3590-3597, 1969  
A:Title: Studies on the active center of human plasmin. Partial amino acid sequence of a  
A:Reference number: A92048; MUID:69234739; PMID:4240117  
A:Contents: annotation: active site  
R:Trexler, M.; Vali, Z.; Pathy, L.  
J. Biol. Chem. 257, 7401-7406, 1982  
A:Title: Structure of the omega-aminocarboxylic acid-binding sites of human plasminogen.  
A:Reference number: A92382; MUID:82213905; PMID:6919539  
A:Contents: annotation: omega-aminocarboxylic acid binding sites  
R:Vali, Z.; Pathy, L.  
J. Biol. Chem. 259, 13690-13694, 1984  
A:Title: The fibrin-binding site of human plasminogen. Arginines 32 and 34 are essential  
A:Reference number: A92458; MUID:85054794; PMID:6094526  
A:Contents: annotation: fibrin binding site; omega-aminocarboxylic acid binding site  
R:Cao, Y.; Ji, R.W.; Davidson, D.; Schaller, J.; Marti, D.; Soehnle, S.; McCance, S.G.;  
J. Biol. Chem. 271, 29461-29467, 1996  
A:Title: Kringle domains of human angiotensin. Characterization of the anti-proliferative  
A:Reference number: A58811; MUID:97067211; PMID:8910613  
A:Contents: annotation  
R:Lijnen, H.R.; Ugwu, F.; Bini, A.; Collen, D.  
Biochemistry 37, 4699-4702, 1998  
A:Title: Generation of an angiotensin-like fragment from plasminogen by stromelysin-1  
A:Reference number: A58812; MUID:9548733; PMID:9548733  
A:Contents: annotation  
R:Tulinsky, A.; Mulichak, A.M.  
submitted to the Brookhaven Protein Data Bank, July 1991  
A:Reference number: A51341; PDB:1PK4  
A:Contents: annotation; X-ray crystallography, 1.9 angstroms, residues 376-454  
R:Tulinsky, A.; Wu, T.P.  
submitted to the Brookhaven Protein Data Bank, July 1991  
A:Reference number: A51488; PDB:2PK4  
A:Contents: annotation; X-ray crystallography, 2.25 angstroms, residues 375-454  
R:Wu, T.P.; Tulinsky, A.  
submitted to the Brookhaven Protein Data Bank, August 1993  
A:Reference number: A51911; PDB:1PKR  
A:Contents: annotation; X-ray crystallography, 2.48 angstroms, residues 102-181  
R:Padmanabhan, K.; Tulinsky, A.  
submitted to the Brookhaven Protein Data Bank, April 1994  
A:Reference number: A52408; PDB:1PMK  
A:Contents: annotation; X-ray crystallography, 2.25 angstroms, residues 377-454  
R:Tulinsky, A.; Mathews, I.I.  
submitted to the Brookhaven Protein Data Bank, December 1995  
A:Reference number: A65244; PDB:1CEA  
A:Contents: annotation; X-ray crystallography, 2.1 angstroms, residues 102-181  
R:Tulinsky, A.; Mathews, I.I.  
submitted to the Brookhaven Protein Data Bank, December 1995  
A:Reference number: A65245; PDB:1CEB  
A:Contents: annotation; X-ray crystallography, 2.1 angstroms, residues 102-181  
R:Mulichak, A.M.; Tulinsky, A.; Ravichandran, K.G.  
Biochemistry 30, 10576-10588, 1991  
A:Title: Crystal and molecular structure of human plasminogen kringle 4 refined at 1.  
A:Reference number: A58819; MUID:92031502; PMID:1657148  
A:Contents: annotation  
R:Wu, T.P.; Padmanabhan, K.; Tulinsky, A.; Mulichak, A.M.  
Biochemistry 30, 10589-10594, 1991  
A:Title: The refined structure of the epsilon-aminocaproic acid complex of human plasminogen  
A:Reference number: A58818; MUID:92031503; PMID:1657149  
A:Contents: annotation  
R:de Vos, A.M.; Uitsch, M.H.; Kelley, R.F.; Padmanabhan, K.; Tulinsky, A.; Westbrook, J.  
Biochemistry 31, 270-279, 1992  
A:Title: Crystal structure of the kringle 2 domain of tissue plasminogen activator at  
A:Reference number: A39483; MUID:92118803; PMID:1310033  
A:Contents: annotation; X-ray crystallography, 2.4 angstroms  
R:Stec, B.; Teeter, M.M.; Whitlow, M.; Yamano, A.  
submitted to the Brookhaven Protein Data Bank, June 1995  
A:Reference number: A65980; PDB:1KRN  
A:Contents: annotation; X-ray crystallography, 1.67 angstroms, residues 376-454  
R:Rejante, M.; Llinas, M.  
submitted to the Brookhaven Protein Data Bank, August 1996  
A:Reference number: A65803; PDB:1LHFJ  
A:Contents: annotation; conformation by (1)H-NMR, residues 103-181  
R:Rejante, M.; Llinas, M.  
submitted to the Brookhaven Protein Data Bank, August 1996  
A:Reference number: A65804; PDB:1LHPK  
A:Contents: annotation; conformation by (1)H-NMR, residues 103-181  
R:Rejante, M.R.; Llinas, M.  
Eur. J. Biochem. 221, 927-937, 1994  
A:Title: (1)H-NMR assignments and secondary structure of human plasminogen kringle 1.  
A:Reference number: S43645; MUID:94237157; PMID:8181475  
A:Contents: annotation; conformation by (1)H-NMR  
R:Rejante, M.R.; Llinas, M.  
Eur. J. Biochem. 221, 939-949, 1994  
A:Title: Solution structure of the epsilon-aminohexanoic acid complex of human plasminogen  
A:Reference number: A58817; MUID:94237158; PMID:8181476  
A:Contents: annotation; conformation by (1)H-NMR  
C:Comment: plasminogen is synthesized by the kidney and is present in plasma and many  
C:Comment: plasminogen is converted to plasmin by plasminogen activators (see PIR:UKH  
d PIR:FGHUGB).



Search completed: April 25, 2003, 15:52:36  
Job time : 51 secs

```
Db 158 YCRNPDNDENGFWCYTTPDKRKYDYCDIPEC-EDKCMHCSGENYEGKIATKMTSGRDCQAW 216
Qy 205 DHQPHRHKFLPERYPDGFDDNYCRNPDGQPRWCYTLDPHTRWEYCAIKTCADNTWMD 264
Db 217 DSQSPHAGYIPSPFNKLNKMYCRNPDGEPFWCFTTDQKRWEFCDIPECT----- 270
Qy 265 TDVPLE--TTTCIOGOGEGYGTWNTIWNIGPCQWDSQYPHEHDMTPENPKCKDLRENY 322
Db 271 TPPSSGPKYQCLGKTGNYGCTVAVTESGHTCQWSEQTPHKHNRTPENFPCKNLENY 330
Qy 323 CRNPDGSESPWCFETDPIRKYCSQIPNCDMSH-----GQDCYRCNG 365
Db 331 CRNPNGERAPWCYTINSVRWEYCT-IPSCSSPLSTERMDVPVPEQTPVPQDCYHNG 389
Qy 366 KNYMGNLSQTRSGLTCSMDKNMEDLHRHFEW-----PDASKLNENYCRNPDDDAHGFWCY 422
Db 390 QSYRGTSTTITGRKCSQSSMTP--HRHLKTPENYPNAG-LTMNYCRNPDAK-KSPWCY 445
Qy 423 TGNPLIPWDYCPISRC 438
Db 446 TTDPRVRWEFCNLKKC 461

RESULT.15
Tl8518
apolipoprotein(a) - western European hedgehog (fragment)
C:Species: Erinaceus europaeus (western European hedgehog)
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 07-Dec-1999
C:Accession: Tl8518
R:Lawn, R.M.; Boonmark, N.W.; Schwartz, K.; Lindahl, G.E.; Wade, D.P.; Byrne, C.D.; Fong
J. Biol. Chem. 270, 24004-24009, 1995
A:Title: The recurring evolution of Lp(a): Insights from cloning of hedgehog apolipoprotein
A:Reference number: 146259; MUID:96025778; PMID:7592597
A:Accession: Tl8518
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-2869 <LAW>
A:Cross-references: EMBL:U33170; NID:gl046358; PID:gl046359; PIDN:AAC48522.1
A:Experimental source: liver
C:Comment: The lipoprotein Lp(a), a major inherited risk factor for atherosclerosis, contains
ent apolipoprotein(a).

Query Match 36.2%; Score 943.5; DB 2; Length 2869;
*Best Local Similarity 45.6%; Pred. No. 6.4e-59;
Matches 173; Conservative 51; Mismatches 126; Indels 29; Gaps 10;

Qy 89 ENKDYIRNCIIGKRSYKGTVSITKSGIKCOPWSSMIPHSFPLSPYRGKDLQENYCRN 148
Db 2490 EPLDLVDQCLEGTGENYRGNAVTAAGHTCQWRQSPHSHSRTPENYPTKNLVGNYCRN 2549
Qy 149 PRGEEGPPWCFTSNPEVRYEVCIDIQCSE-----VECMTCNGESYRGLMDHTES 197
Db 2550 PDGEV-APWCYTNSAMRWYCISIPACESPTPTTEHLVPEQCLEGGNGENYQGNNAITVS 2608
Qy 198 GKICORWDHQTPIRHKFLPERYPDKGFDNDYCRNPDGQPRWCYTLDPHTRWEYCAIKTC 257
Db 2609 GQPCQGWKQKTPHREYTPENYPSKNLFNGYCRNPDGAIAPWCYTNTSAVRWEYCSIPTC 2668
Qy 258 ADNTMNDTDVPLETTECIQGGEGYRGVTWNTIWNIGPCQWDSQYPHEHDMTPENPKCKD 317
Db 2669 -ESSSPTEPMIIPDQCLEGTGENYRGNAVTAAGHTCQWRQSPHSHSRTPENYPTKN 2727
Qy 318 LRENYCRNPDGSESPWCFETDPIRKYCSQIPNCDMSH-----GQD-CYRGNKKNYM 369
Db 2728 LFGNYCRNPDGEVAPWCYTNSAVRWYC-KIPSCNSSSPTEPSTSDQCLKNGESYQ 2786
Qy 370 GNLSQTRSGLTCSMDKNMEDLHRHFEWPD---ASKLNENYCRNPDDDAHGFWCYTGNP 426
Db 2787 GNISVTSGYTCQWRRE--QTPHRA-RTPDNYPCKNLVGNYCRNPDGEV-APWCYTNS 2842
Qy 427 LIPWDYCPISRCGDTTPT 445
Db 2843 AVRWEYCSIPTCESPTPT 2861
```

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.4.p5.4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 25, 2003, 15:45:29 ; Search time 25 Seconds  
(without alignments)  
741.597 Million cell updates/sec

Title: US-09-674-377b-1

Perfect score: 2604

Sequence: 1 ERRRRNTIHEFKSAKTTLI.....IPWDYCPISRCEGDTTPTIV 447

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_40.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	2601	99.9	728	1	HGF_HUMAN
2	2408	92.5	728	1	HGF_MOUSE
3	2401	92.2	728	1	HGF_RAT
4	1180.5	45.3	711	1	HGFL_HUMAN
5	1125	43.2	716	1	HGFL_MOUSE
6	1066	40.9	790	1	PLMN_PIG
7	1059	40.7	812	1	PLMN_MOUSE
8	1053.5	40.5	810	1	PLMN_ERIEU
9	1044.5	40.1	810	1	PLMN_MACMU
10	1042	40.0	810	1	PLMN_HUMAN
11	1038	39.9	812	1	PLMN_BOVIN
12	879	33.8	4548	1	APOA_HUMAN
13	843.5	32.4	1420	1	APOA_MACMU
14	408.5	15.7	325	1	PLMN_PETMA
15	357.5	13.7	625	1	THRB_BOVIN
16	353	13.6	169	1	THRB_MOUSE
17	352.5	13.5	618	1	THRB_MOUSE
18	350	13.4	622	1	THRB_HUMAN
19	344	13.2	617	1	THRB_RAT
20	307.5	11.8	559	1	TPA_RAT
21	303	11.6	562	1	TPA_HUMAN
22	303	11.6	566	1	TPA_BOVIN
23	297.5	11.4	559	1	TPA_MOUSE
24	253	9.7	338	1	PLMN_HORSE
25	241	9.3	343	1	PLMN_SHEEP
26	237	9.1	333	1	PLMN_CANFA
27	230	8.8	943	1	ROR2_HUMAN
28	228.5	8.8	603	1	FA12_CAVPO
29	223	8.6	944	1	ROR2_MOUSE
30	220	8.4	937	1	ROR1_MOUSE
31	217	8.3	937	1	ROR1_HUMAN
32	204	7.8	655	1	HGFA_HUMAN
33	198	7.6	653	1	HGFA_MOUSE

RESULT 1  
HGF\_HUMAN  
ID HGF\_HUMAN STANDARD; PRT; 728 AA.  
AC P14210; Q9UDU6; Q9BYL9;  
DT 01-JAN-1990 (Rel. 13, Created)  
DT 01-AUG-1991 (Rel. 19, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Hepatocyte growth factor precursor (Scatter factor) (SF)  
DE (Hepatopoietin A).  
GN HGF OR HPTA.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=91340155; PubMed=1831432;  
RA Seki T., Hagiya M., Shimonishi M., Nakamura T., Shimizu S.;  
RT "Organization of the human hepatocyte growth factor-encoding gene.";  
RL Gene 102:213-219(1991).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX TISSUE=Placenta;  
RA MEDLINE=89392017; PubMed=2528952;  
RA Miyazawa K., Tsubouchi H., Naka D., Takahashi K., Okigaki M.,  
RA Arakaki N., Nakayama H., Hirono S., Sakiyama O., Takahashi K.,  
RA Gohda E., Daikuhara Y., Kitamura N.;  
RT "Molecular cloning and sequence analysis of cDNA for human hepatocyte  
growth factor.";  
RL Biochem. Biophys. Res. Commun. 163:967-973(1989).  
RN [3]  
RP SEQUENCE FROM N.A.  
RX TISSUE=Leukocyte;  
RA MEDLINE=91025062; PubMed=2145836;  
RA Seki T., Ihara I., Sugimura A., Shimonishi M., Nishizawa T.,  
RA Asami O., Hagiya M., Nakamura T., Shimizu S.;  
RT "Isolation and expression of cDNA for different forms of hepatocyte  
growth factor from human leukocyte.";  
RL Biochem. Biophys. Res. Commun. 172:321-327(1990).  
RN [4]  
RP SEQUENCE FROM N.A., AND SEQUENCE OF 55-73 AND 495-520.  
RX TISSUE=Liver;  
RA MEDLINE=90066676; PubMed=2531289;  
RA Nakamura T., Nishizawa T., Hagiya M., Seki T., Shimonishi M.,  
RA Sugimura A., Tashiro K., Shimizu S.;  
RT "Molecular cloning and expression of human hepatocyte growth factor.";  
RL Nature 342:440-443(1989).  
RN [5]  
RP SEQUENCE FROM N.A.  
RX TISSUE=Embryonic fibroblast;  
RA MEDLINE=91334393; PubMed=1831266;  
RA Weidner K.M., Arakaki N., Hartmann G., Vandekerckhove J., Weingart S.,  
RA Rieder H., Fonatsch C., Tsubouchi H., Hishida T., Daikuhara Y.,  
RA Birchnieker W.;  
RT "Evidence for the identity of human scatter factor and human  
hepatocyte growth factor.";

34 195 7.5 615 1 FA12\_HUMAN p00748 homo sapien  
35 189 7.3 593 1 FA12\_BOVIN p98140 bos taurus  
36 187 7.2 473 1 KREM\_MOUSE Q99n43 mus musculus  
37 187 7.2 473 1 KREM\_RAT Q92484 rattus norv  
38 180.5 6.9 475 1 KREM\_HUMAN Q96mu8 homo sapien  
39 158.5 6.1 433 1 UROK\_BOVIN Q05589 bos taurus  
40 158 6.1 442 1 UROK\_PIG P04185 sus scrofa  
41 156 6.0 477 1 URT1\_DESRO p98119 desmodus ro  
42 152.5 5.9 875 1 NETR\_HUMAN p56730 homo sapien  
43 151.5 5.8 431 1 UROK\_HUMAN p00749 homo sapien  
44 151 5.8 432 1 UROK\_RAT p29598 rattus norv  
45 151 5.8 477 1 URT2\_DESRO p15638 desmodus ro

Proc. Natl. Acad. Sci. U.S.A. 88:7001-7005(1991).

[6] SEQUENCE FROM N.A.  
Courtney L., Elliot G., Angell S.;  
Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.

[7] SEQUENCE OF 249-695 FROM N.A.  
MEDLINE-9136928; PubMed-1832556;  
MEDLINE-9136928; PubMed-1832556;  
Miyazawa K., Kitamura A., Kitamura N.;  
"Structural organization and the transcription initiation site of the  
human hepatocyte growth factor gene";  
Biochemistry 30:9170-9176(1991).

[8] SIGNAL SEQUENCE CLEAVAGE SITE.  
MEDLINE-91207365; PubMed-1826837;  
Yoshiyama Y., Arakaki N., Naka D., Takahashi K., Hirono S., Kondo J.,  
Kobayashi T., Sakai H., Inoue H., Ishii T.,  
Nakayama H., Gohta E., Kitamura N., Tsubouchi H.;  
Hishida T., Daikuhara Y.;  
"Identification of the N-terminal residue of the heavy chain of both  
native and recombinant human hepatocyte growth factor";  
Biochem. Biophys. Res. Commun. 175:660-667(1991).

[9] CARBOHYDRATE-LINKAGE SITE THR-476.  
MEDLINE-93129192; PubMed-1482348;  
Shimizu N., Hara H., Sogabe T., Sakai H., Ihara I., Inoue H.,  
Nakamura T., Shimizu S.;  
"Hepatocyte growth factor is linked by O-glycosylated oligosaccharide  
on the alpha chain.";  
Biochem. Biophys. Res. Commun. 189:1329-1335(1992).

[10] MUTAGENESIS.  
MEDLINE-92331602; PubMed-1321034;  
Lokker N.A., Mark M.R., Luis E.A., Bennett G.L., Robbins K.A.,  
Baker J.B., Godowski P.J.;  
"Structure-function analysis of hepatocyte growth factor:  
identification of variants that lack mitogenic activity yet retain  
high affinity receptor binding.";  
EMBO J. 11:2503-2510(1992).

[11] STRUCTURE BY NMR OF 31-127.  
MEDLINE-98154323; PubMed-9493272;  
Zhou H., Mazzulla M.J., Kaufman J.D., Stahl S.J., Wingfield P.T.,  
Rubin J.S., Bottaro D.P., Byrd R.A.;  
"The solution structure of the N-terminal domain of hepatocyte growth  
factor reveals a potential heparin-binding site.";  
Structure 6:109-116(1998).

[12] X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF 35-210.  
MEDLINE-99036858; PubMed-9817840;  
Ullsch M., Lokker N.A., Godowski P.J., de Vos A.M.;  
"Crystal structure of the NK1 fragment of human hepatocyte growth  
factor at 2.0-A resolution.";  
Structure 6:1383-1393(1998).  
!- FUNCTION: HGF IS A POTENT MITOGEN FOR MATURE PARENCHYMAL  
HEPATOCTE CELLS, SEEMS TO BE AN HEPATOTROPIC FACTOR, AND ACTS  
AS GROWTH FACTOR FOR A BROAD SPECTRUM OF TISSUES AND CELL TYPES.  
IT HAS NO DETECTABLE PROTEASE ACTIVITY.  
!- SUBUNIT: DIMER OF AN ALPHA CHAIN AND A BETA CHAIN LINKED BY A  
DISULFIDE BOND.  
!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1. PLASMINOGEN SUBFAMILY.  
!- SIMILARITY: CONTAINS 4 KRINGLE DOMAINS.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).

EMBL: D90334; BAAL4348.1;  
EMBL: D90318; BAAL4348.1; JOINED.  
EMBL: D90319; BAAL4348.1; JOINED.

Query Match	99.98;	Score 2601;	DB 1;	Length 728;
Best Local Similarity	99.84;	Pred. No. 8.3e-192;		
Matches 446;	Conservative 1;	Mismatches 0;	Indels 0;	Gaps 0;
Qy 1	ERKRRNTIHEFKSAKTLIKIDPALKIKTKVNTADOCANCRCTNKGKLPETCKAFVFDK	60		
Db 32	QKRRNTIHEFKSAKTLIKIDPALKIKTKVNTADOCANCRCTNKGKLPETCKAFVFDK	91		
Qy 61	ARKQCLWFPFNSMSGVKKFEGHEFDLYENKDYIRNCIIIGKRSYKGVTSITKSGIKQCP	120		
Db 92	ARKQCLWFPFNSMSGVKKFEGHEFDLYENKDYIRNCIIIGKRSYKGVTSITKSGIKQCP	151		
Qy 121	WSSMIPHEHSELPSSYRGKDLQENYCRNPRGEEGPMCFSTNPEVRYEVCIDIPOCSEVEC	180		
Db 152	WSSMIPHEHSELPSSYRGKDLQENYCRNPRGEEGPMCFSTNPEVRYEVCIDIPOCSEVEC	211		
Qy 181	MTNCGESYRGLMDHTESGKICQRDHQPFRHKKFLPERYPDKGDDNCRNPDGQPRWC	240		
Db 212	MTNCGESYRGLMDHTESGKICQRDHQPFRHKKFLPERYPDKGDDNCRNPDGQPRWC	271		
Qy 241	YTLDPHTRWECYCAKTCADNTMNDTVDPLETTECIQOGEGYRGVNTIWNIGPCQRWDS	300		
Db 272	YTLDPHTRWECYCAKTCADNTMNDTVDPLETTECIQOGEGYRGVNTIWNIGPCQRWDS	331		
Qy 301	QYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDMSHGQDC	360		
Db 332	QYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDMSHGQDC	391		
Qy 361	YRGNGKNTYMGNSQTRSGLTCSMDKNNMEDLHRHFWEPDASKLNNYCRNPDHAGPW	420		
Db 392	YRGNGKNTYMGNSQTRSGLTCSMDKNNMEDLHRHFWEPDASKLNNYCRNPDHAGPW	451		
Qy 421	CYTGNPLIPWDYCPISRCGEGDTTPIV 447			
Db 452	CYTGNPLIPWDYCPISRCGEGDTTPIV 478			
RESULT 2				
ID	HGF_MOUSE	STANDARD;	PRF;	728 AA.
AC	008048; 064007; 061662;			
DT	01-NOV-1995 (Rel. 32, Created)			
DT	01-NOV-1995 (Rel. 32, Last sequence update)			
DT	15-JUN-2002 (Rel. 41, Last annotation update)			
DE	Hepatocyte growth factor precursor (Scatter factor) (SF)			
DE	(Hepatopoietin A).			
GN	HGF.			
OS	Mus musculus (Mouse).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
OX	NCBI_TaxID=10090;			
RN	[1]			
RP	SEQUENCE FROM N.A. (LONG AND SHORT ISOFORMS), AND SEQUENCE OF 496-504.			
RC	TISSUE=Mammary fibroblast;			
RX	MEDLINE=94183257; PubMed=8135822;			
RA	Sasaki M., Nishio M., Sasaki T., Enami J.;			
RT	"Identification of mouse mammary fibroblast-derived mammary growth			
RT	factor as hepatocyte growth factor.";			
RL	Biochem. Biophys. Res. Commun. 199:772-779(1994).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Liver;			
RX	MEDLINE=94363381; PubMed=8081873;			
RA	Lee C.C., Kozak C.A., Yamada K.M.;			
RT	"Structure, genetic mapping, and expression of the mouse Hgf/scatter			
RT	factor gene.";			
RL	Cell Adhes. Commun. 1:101-111(1993).			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Liver;			
RX	MEDLINE=94060105; PubMed=8241272;			
RA	Liu Y., Michalopoulos G.K., Zarnegar R.;			

RT	"Molecular cloning and characterization of cDNA encoding mouse			
RT	hepatocyte growth factor.";			
RL	Biochim. Biophys. Acta 1216:299-303(1993).			
CC	-!- FUNCTION: HGF IS A POTENT MITOGEN FOR MATURE PARENCHYMAL			
CC	HEPATOCYTE CELLS, SEEMS TO BE AN HEPATOTROPIC FACTOR, AND ACTS			
CC	AS GROWTH FACTOR FOR A BROAD SPECTRUM OF TISSUES AND CELL TYPES.			
CC	IT HAS NO DETECTABLE PROTEASE ACTIVITY.			
CC	-!- SUBUNIT: DIMER OF AN ALPHA CHAIN AND A BETA CHAIN LINKED BY A			
CC	DISULFIDE BOND.			
CC	-!- ALTERNATIVE PRODUCTS: 2 isoforms; a long form (shown here) and a			
CC	short form; are produced by alternative splicing.			
CC	-!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1. PLASMINOGEN SUBFAMILY.			
CC	-!- SIMILARITY: CONTAINS 4 KRINGLE DOMAINS.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/			
CC	or send an email to license@isb-sib.ch).			
CC	-----			
DR	EMBL; D10212; BAA01064.1; -			
DR	EMBL; D10213; BAA01065.1; -			
DR	EMBL; S71816; AAB31855.1; -			
DR	EMBL; X72307; CAA51054.1; ALT_INIT.			
DR	HSSP; P14210; 1BHT.			
DR	MEROPS; S01.982; -			
DR	MGD; MGI:96079; Hgf.			
DR	InterPro; IPR001314; Chymotrypsin.			
DR	InterPro; IPR000001; Kringle.			
DR	InterPro; IPR003014; PAN.			
DR	InterPro; IPR003609; Pan_app.			
DR	InterPro; IPR001254; Ser_protease_Try.			
DR	Pfam; PF00024; PAN. 1.			
DR	Pfam; PF00051; kringle; 4.			
DR	Pfam; PF00089; trypsin; 1.			
DR	PRINTS; PR00722; CHYMOTRYPSIN.			
DR	PRINTS; PR00018; KRINGLE.			
DR	ProDom; PD000395; Kringle; 4.			
DR	SMART; SM00130; KR; 4.			
DR	SMART; SM00473; PAN_AP; 1.			
DR	SMART; SM00020; Tryp_Spc; 1.			
DR	PROSITE; PS00021; KRINGLE_1; 4.			
DR	PROSITE; PS00070; KRINGLE_2; 4.			
DR	PROSITE; PS0240; TRYPSIN_DOM; 1.			
KW	Growth factor; Kringle; Glycoprotein; Serine protease homolog;			
KW	Repeat; Signal; Alternative splicing.			
FT	SIGNAL 1 32			
FT	CHAIN 33 495			
FT	CHAIN 496 728			
FT	MOD_RES 33 33			
FT	-----			
FT	DOMAIN 33 128			
FT	DOMAIN 129 207			
FT	DOMAIN 212 289			
FT	DOMAIN 306 384			
FT	DOMAIN 392 470			
FT	DOMAIN 496 728			
FT	DISULFID 71 97			
FT	DISULFID 75 85			
FT	DISULFID 488 607			
FT	CARBOHYD 295 295			
FT	CARBOHYD 403 403			
FT	CARBOHYD 569 569			
FT	CARBOHYD 656 656			
FT	VARSPIC 163 167			
FT	CONFLICT 344 344			
FT	CONFLICT 479 479			
FT	CONFLICT 564 564			
SQ	SEQUENCE 728 AA; 82944 MW; A0381FC497534328 CRC64;			

Query Match

92.5%; Score 2408; DB 1; Length 728;





Db 213 MTCNGSYRGPMDHTEGKTCQRMDOQTPIRHKFLPERYPDKGFDNDYCRNPDGKPRPW 272  
QY 241 YTLDPHTRWYCAIKTCADTMDTNDVPLETTICOGEGYRGTVNTWNGIPCORWDS 300  
Db 273 YTLDPDPWEYCAIKWCAHSAVNETDVPMTTECIKOGEGYRGTVNTWNGIPCORWDS 332  
QY 301 QYPHEHMTPEFKCKDLRENYCRNPDGSSPMCFDTPDNIIRVGYCSQIPNCDMSHGQDC 360  
Db 333 QYPKHKHDITPEFKCKDLRENYCRNPDGSAESPMCFDTPDNIIRVGYCSQIPKCDVSSGDC 392  
QY 361 YRGNKKNYGNLSQTRSGLTCSMWKKNMEDLHRHIFWEPDASKNLNENYCRNPDHAGPW 420  
Db 393 YRGNKKNYGNLSKTRSGLTCSMWKKNMEDLHRHIFWEPDASKLTKNYCRNPDHAGPW 452  
QY 421 CYTGNPLIPWDYCPISRCCEGDTPTIV 447  
Db 453 CYTGNPLVWDYCPISRCCEGDTPTIV 479  
RESULT 4  
HGFL\_HUMAN STANDARD; PRT; 711 AA.  
AC P26927: Q13350; Q14870;  
DT 01-AUG-1992 (Rel. 23, Created)  
DT 01-AUG-1992 (Rel. 23, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Hepatocyte growth factor-like protein precursor (Macrophage  
stimulatory protein) (MSP) (Macrophage stimulating protein).  
GN MST1 OR HGFL.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Liver;  
RX MEDLINE=92002016; PubMed=1655021;  
RA Han S., Stuart L.A., Frieznier Degen S.J.;  
RT "Characterization of the DNFI52 locus on human chromosome 3:  
identification of a gene coding for four kringle domains with  
homology to hepatocyte growth factor.";  
RL Biochemistry 30:9768-9780(1991).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Liver;  
RX MEDLINE=93340141; PubMed=8393443;  
RA Yoshimura T., Yuhki N., Wang M.H., Skeel A., Leonard E.J.;  
RT "Cloning, sequencing, and expression of human macrophage stimulating  
protein (MSP, MST1) confirms MSP as a member of the family of kringle  
proteins and locates the MSP gene on chromosome 3.";  
RL J. Biol. Chem. 268:15461-15468(1993).  
CC -1- FUNCTION: PROBABLY HAS NO PROTEOLYTIC ACTIVITY, SINCE CRUCIAL AA  
CHARACTERISTIC OF SERINE PROTEASES CATALYTIC SITES ARE NOT  
CONSERVED.  
CC -1- PTM: MAY BE CLEAVED AFTER AA 484, TO YIELD A TWO-CHAIN MOLECULE  
HELLED TOGETHER BY DISULFIDE BONDS, OR TWO SEPARATE POLYPEPTIDES.  
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1. PLASMINOGEN SUBFAMILY.  
CC -1- SIMILARITY: CONTAINS 4 KRINGLE DOMAINS.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; M74178; AAC50165.1; -  
CC EMBL; U37055; AAC50471.1; -  
CC EMBL; L11924; AAC59872.1; -  
CC PIR; A40331; A40331.  
CC HSP; P00747; 2PK4.

DR MEROPS; S01.975; -  
DR Genew; HGNC:7380; MST1.  
DR MIM; 142408;  
DR InterPro; IPR001314; Chymotrypsin.  
DR InterPro; IPR000001; Kringle.  
DR InterPro; IPR003014; PAN.  
DR InterPro; IPR003609; Pan\_app.  
DR InterPro; IPR001254; Ser\_protease\_Try.  
DR Pfam; PF00051; kringle; 4.  
DR Pfam; PF00089; trypsin; 1.  
DR PRINTS; PR00722; CHYMOTRYPSIN.  
DR PRINTS; PR00018; KRINGLE.  
DR PRODOM; PD000395; Kringle; 4.  
DR SMART; SM00130; KR; 4.  
DR SMART; SM00473; PAN\_AP; 1.  
DR SMART; SM00020; TRYD\_SPC; 1.  
DR PROSITE; PS00021; KRINGLE\_1; 4.  
DR PROSITE; PS00070; KRINGLE\_2; 4.  
DR PROSITE; PS00240; TRYPSIN\_DOM; 1.  
KW Kringle; Glycoprotein; Serine protease homolog; Repeat; Signal;  
KW Polymorphism.  
FT SIGNAL 1 18 POTENTIAL.  
FT CHAIN 19 711 HEPATOCYTE GROWTH FACTOR-LIKE PROTEIN.  
FT DOMAIN 32 109 PAP.  
FT DOMAIN 110 186 KRINGLE 1.  
FT DOMAIN 191 268 KRINGLE 2.  
FT DOMAIN 283 361 KRINGLE 3.  
FT DOMAIN 370 448 KRINGLE 4.  
FT DOMAIN 484 711 SERINE PROTEASE-LIKE.  
FT DISULFID 56 78 BY SIMILARITY.  
FT DISULFID 60 66 BY SIMILARITY.  
FT DISULFID 110 186 BY SIMILARITY.  
FT DISULFID 131 169 BY SIMILARITY.  
FT DISULFID 157 181 BY SIMILARITY.  
FT DISULFID 191 268 BY SIMILARITY.  
FT DISULFID 194 324 INTERCHAIN (BY SIMILARITY).  
FT DISULFID 212 251 BY SIMILARITY.  
FT DISULFID 240 263 BY SIMILARITY.  
FT DISULFID 283 361 BY SIMILARITY.  
FT DISULFID 304 343 BY SIMILARITY.  
FT DISULFID 332 355 BY SIMILARITY.  
FT DISULFID 370 448 BY SIMILARITY.  
FT DISULFID 391 431 BY SIMILARITY.  
FT DISULFID 419 443 INTERCHAIN (BY SIMILARITY).  
FT DISULFID 468 588 BY SIMILARITY.  
FT DISULFID 507 523 BY SIMILARITY.  
FT DISULFID 602 667 BY SIMILARITY.  
FT DISULFID 632 646 BY SIMILARITY.  
FT DISULFID 657 685 BY SIMILARITY.  
FT CARBOHYD 72 72 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 296 296 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 615 615 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT VARIANT 13 13 Y -> C.  
FT VARIANT 212 212 /FTID=VAR\_006631.  
FT VARIANT 623 623 C -> F.  
FT CONFLICT 623 623 /FTID=VAR\_006632.  
FT CONFLICT 711 AA; 80379 MW; 596ED21F180290E4 CRC64;  
SQ SEQUENCE 711 AA; 80379 MW; 596ED21F180290E4 CRC64;  
Query Match 45.3%; Score 1180.5; DB 1; Length 711;  
Best Local Similarity 47.4%; Pred. No. 4.4e-83;  
Matches 197; Conservative 62; Mismatches 148; Indels 9; Gaps 4;  
QY 33 VNTADOCANCTRNKGLPFTCKAFVDFKARKQCLWFFNSMSGVKKFEGHEFDLYENKD 92  
Db 50 VADREECAGRC-----GLPMDCRAFHYNVSSHGCCOLLPTQHSPTLRRSRCGLDFOKKD 105  
QY 93 YIRNCIIIGKRSYKGVTSITKSGIKCQWSSMIPHEHSFLPSSYRGKDLQENYCRNPRGE 152  
Db 106 YVRTCINNGVGYRGTMATVGGGLPCQAWSHKFPNDHKYPTPLNG--LEENFCRNPDGD 163  
QY 153 EGGPWCTSNPEVRYEYCDIPQCSEVECMTCNGESYRGLMDHTESGKICQRWDHQTPIRH 212

```
Db 164 PGGWCYTDDPAVFQSGGKSCREACWCNGEYRGAVDRTESGRCQWDLQHPQH 223
Qy 213 KFLPERYPDKGFDNDYCRNPDGQPRPCWYTLDPHTRWEYCAIKTCADTWNDDVPLETT 272
Db 224 PFEFGKFLDQGLDDNYCRNPDGSRPCWYTTDPOIEREFCDLPRGSEAPQGEA--TTV 281
Qy 273 ECIQOGEGYRGVTWNTIWNIGTPCQWDSQYVPEHDMTPENFKCKDLRENCRNPDGSESP 332
Db 282 SCFRKGEGYRGVTANTTAGVPCQWDAQIPHQHRTPEKYACKDLRENCRNPDGSEAP 341
Qy 333 WCFTDPNIRVGYCSQIPNC--DMSHGQDCYRGNGKNTMGNLSQTRSLGTCMMWKNMEDL 391
Db 342 WCFTLPCRMRAFCYQIRRCTDDVRPQDCYHAGEQYRGVTSKTRKGVCQCRWSAETPHK 401
Qy 392 HRHIFWEPDASKLENYCRNPDGDAHGPWCYTGPNLPLPDWYCPISRCGDTPTIV 447
Db 402 PQFTFSPHAQLEENFCRNPDSHGSPWCYTMDPRTPFYCALRRCAADDQPPSIL 457

RESULT 5
HGFL_MOUSE STANDARD; PRT; 716 AA.
AC P26928;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DE 15-JUN-2002 (Rel. 41, Last annotation update)
DE Hepatocyte growth factor-like protein precursor (Macrophage
DE stimulatory protein) (MSP).
GN MST1 OR HGFL.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_taxid=10909;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/c; TISSUE=Liver;
RX MEDLINE=92002017; PubMed=1832957;
RA Friezen-Degen S.J., Stuart L.A., Han S., Jamison C.S.;
RT "Characterization of the mouse cDNA and gene coding for a hepatocyte
RT growth factor-like protein: expression during development.";
RL Biochem J 30:9781-9791(1991).
CC !- FUNCTION: PROBABLY HAS NO PROTEOLYTIC ACTIVITY, SINCE CRUCIAL AA
CC CHARACTERISTIC OF SERINE PROTEASES CATALYTIC SITES ARE NOT
CC CONSERVED.
CC !- TISSUE SPECIFICITY: LIVER. LOWER LEVELS IN LUNG, PLACENTA AND
CC ADRENAL.
CC !- DEVELOPMENTAL STAGE: IS EXPRESSED AT LOW LEVELS DURING GESTATION.
CC STABLE AFTERWARDS.
CC !- PTM: MAY BE CLEAVED AFTER AA 488, TO YIELD A TWO-CHAIN MOLECULE
CC HELD TOGETHER BY DISULFIDE BONDS, OR CLEAVED INTO TWO SEPARATE
CC POLYPEPTIDES.
CC !- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1. PLASMINOGEN SUBFAMILY.
CC !- SIMILARITY: CONTAINS 4 KRINGLE DOMAINS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
```

```
DR InterPro: IPR001254; Ser_protease_Try.
DR Pfam: PF00024; PAN; 1.
DR Pfam: PF00051; kringle; 4.
DR Pfam: PF00089; trypsin; 1.
DR PRINTS: PRO0722; CHYMOTRYPSIN.
DR PRINTS: PRO0018; KRINGLE.
DR ProDom: PD000395; kringle; 4.
DR SMART: SM00130; KR; 4.
DR SMART: SM00473; PANAP; 1.
DR SMART: SM00020; TRYD-SPC; 1.
DR PROSITE: PS00021; KRINGLE_1; 4.
DR PROSITE: PS0070; KRINGLE_2; 4.
DR PROSITE: PS0240; TRYPSIN_DOM; 1.
KW Kringle; Glycoprotein; Serine protease homolog; Repeat; Signal.
FT SIGNAL 1 18
FT CHAIN 19 716
FT DOMAIN 19 109
FT DOMAIN 110 186
FT DOMAIN 191 268
FT DOMAIN 292 370
FT DOMAIN 379 457
FT DOMAIN 489 716
FT DISULFID 56 78
FT DISULFID 60 66
FT DISULFID 110 186
FT DISULFID 131 169
FT DISULFID 157 181
FT DISULFID 191 268
FT DISULFID 194 333
FT DISULFID 212 251
FT DISULFID 240 263
FT DISULFID 292 370
FT DISULFID 313 352
FT DISULFID 341 364
FT DISULFID 379 457
FT DISULFID 400 440
FT DISULFID 428 452
FT DISULFID 477 593
FT DISULFID 512 528
FT DISULFID 607 672
FT DISULFID 637 651
FT DISULFID 662 690
FT CARBOHYD 72 72
FT CARBOHYD 173 173
FT CARBOHYD 305 305
FT CARBOHYD 620 620
FT CONFLICT 19 19
SQ SEQUENCE 716 AA; 80588 MW; BBCE02EF85213ACC CRC64;
Query Match 43.28; Score 1125; DB 1; Length 716;
Best Local Similarity 45.98; Pred. No. 7.9e-79;
Matches 195; Conservative 69; Mismatches 143; Indels 18; Gaps 6;
Qy 33 VNTADQACNRCTRNGKLPFTCKAFVDFKARKQCLWFFNFNSMSGVKKKEFGHEFDLYENKD 92
Db 50 VADAEACARR-----GPLLDCRAFYNNSSHCQQLLPWQHSHTQLYHSSLCHFQKDK 105
Qy 93 YIRNCIIGKGRSYKGTSVITKSGIKQCPWSSMIPHEHSFLPSSYRGKDLQNYCRNPRGE 152
Db 106 YVRTCIMDNGVSYRGTVARTAGGLPCQAWSRRFFNDHKYTPTPKNG--LEENFCRNPDDG 163
Qy 153 EGGPWCFTSNPEVRYEVCDDIPQCSEVCEMTCNGESYRGLMDHTESGKTCQWMDHOTPHR 212
Db 164 PRGPWCYTTNRSVRFQSGIGTKCREAVCVLCNGEDYRGVDVETSGREGQWDLQHPHSH 223
Qy 213 KFLPERYPDKGFDNDYCRNPDGQPRPCWYTLDPHTRWEYCAIKTCADN-----TMDT 265
Db 224 PFEFGKFLDQGLDDNYCRNPDGSRPCWYTTDPOIEREFCDLPRGSEAPLPTTVKGSQR 283
Qy 266 DVPLETTECIQOGEGYRGVTWNTIWNIGTPCQWDSQYVPEHDMTPENFKCKDLRENCRN 325
Db 284 RNKGKALNCFRGKGDYRGTTNTTSAGVPCQWDAQSPHQHFRFVPEKYACKDLRENFCRN 343
```

DR	HSSP: P00747; 5HPG.	DR
DR	MEROPS; S01.233; .	DR
DR	GlycoSuiteDB; p06867; .	DR
DR	InterPro: IPR001314; Chymotrypsin.	DR
DR	InterPro: IPR000001; Kringle.	DR
DR	InterPro: IPR003014; PAN.	DR
DR	InterPro: IPR003609; Pan_app.	DR
DR	InterPro: IPR001254; ser_protease_Try.	DR
DR	Pfam; PF00024; PAN; 1.	DR
DR	Pfam; PF00051; kringle; 5.	DR
DR	Pfam; PF00089; trypsin; 1.	DR
DR	PRINTS; PR00722; CHYMOTRYPSIN.	DR
DR	PRINTS; PR00018; KRINGLE.	DR
DR	ProDom; PD000395; Kringle; 5.	DR
DR	SMART; SM00130; KR; 5.	DR
DR	SMART; SM00473; PAN_AP; 1.	DR
DR	SMART; SM00020; Tryp_SPC; 1.	DR
DR	PROSITE; PS00021; KRINGLE_1; 5.	DR
DR	PROSITE; PS50070; KRINGLE_2; 5.	DR
DR	PROSITE; PS50240; TRYPsin_DOM; 1.	DR
DR	PROSITE; PS00134; TRYPsin_HIS; 1.	DR
DR	PROSITE; PS00135; TRYPsin_SPR; 1.	DR
KW	Hydrolase; Serine protease; Plasma; Glycoprotein; Fibrinolysis;	
KW	Tissue remodeling; Blood coagulation; Kringle; Zymogen; Repeat.	
FT	CHAIN 1 560 PLASMIN HEAVY CHAIN A.	
FT	CHAIN 561 790 SERINE PROTEASE.	
FT	DOMAIN 84 162 KRINGLE 1.	
FT	DOMAIN 166 243 KRINGLE 2.	
FT	DOMAIN 256 333 KRINGLE 3.	
FT	DOMAIN 358 435 KRINGLE 4.	
FT	DOMAIN 461 540 KRINGLE 5.	
FT	ACT_SITE 602 602 CHARGE RELAY SYSTEM.	
FT	ACT_SITE 645 645 CHARGE RELAY SYSTEM.	
FT	ACT_SITE 740 740 CHARGE RELAY SYSTEM.	
FT	CARBOHYD 289 289 N-LINKED (GLCNAC. .).	
FT	CARBOHYD 340 340 /FTIG-CAR_000019.	
FT	CARBOHYD 340 340 /FTIG-CAR_000020.	
FT	SEQUENCE 790 AA; 88592 MW; F04EA06E74BCD58E CRC64;	
QY	Query Match 40.9%; Score 1066; DB 1; Length 790;	
DB	Best Local Similarity 44.6%; Pred. No. 2.9e-74;	
DB	Matches 195; Conservative 65; Mismatches 137; Indels 40; Gaps	
QY	25 ALKIKTKVNTAQCARNCTRNGKLPTCKAFVDRKARKOCLMFPNSMSSGVKKSEFGE 84	
DB	16 SLRSKQVAARSVEECAAKCAETN--FICRAFOYHSKQOCQVMAENSKTSPARM--RD 71	
QY	85 FDLYENKDYIRNIGIKGRYSKGTIVTSIKGKICQCPWSSMIPHEHSLFPSSYRGKDLQEN 144	
DB	72 VLFPEKRIYLSECKTGNGKYNRGTTSKTSKSVICQKWSVSSPHIKYSPKFPFLAGLEEN 131	
QY	145 YCRNPRGEEGGPWCFTNSPEVREYVCDIPQCSVEVCMTNGESYRGLMDHTESGKTCQRW 204	
DB	132 YCRNPNNDERGKPCWYTTDPETRFDYCDIPCC-EDECMHCSGEHYEGKISKTMGIECQSW 190	
QY	205 DHQTPRHKKFLPERYDPKGFDNNYCRNPDCGQPRWCYTLDPHTRWEVCAIKTCADNTMND 264	
DB	191 GSQSPHAGHYLPKSPFNKMLKMYCRNPDCGEPKRWCFITDPNKRWFCFDIPRC--TTPPP 248	
QY	265 TDVPLETTECICQGGEGYRGTVNTIWNIGICPQRWDSOY PHEHDMTIPENFKCKDLRENYCR 324	
DB	249 TSGP--TYQCLKRGENYRGTVSVTASGHTCQRNQAQSPHKHNRTPENFPCKLNEENYCR 306	
QY	325 NPDGSESPWCFTTDPNIRVGYCSQIPLNCDS-----H-----GQDCYRGNKN 367	
DB	307 NPDGETAPWCYTTDSEVRWDYC-KIPSCGSSTTSTEHLDAVPPEQTPVAQDCYRGNES 365	
QY	368 YMGNLSTRGLSCWMDKKNWDLHRRHIFWE-----PDASKLNEYCRNPDDDAGHPWC 421	
DB	366 YRGTSSTTITGRKCSQSVSWTS--HRR----EKTGPNFPNAG-LTMNYCRNPAD-KSPWC 418	

QY 422 YTCNPLIPWDYCPISRC 438  
DB 419 YTTDPVRWEYCNKKC 435

## RESULT 7

```

PLMN_MOUSE
ID PLMN_MOUSE STANDARD; PRT; 812 AA.
AC P20918;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Plasminogen precursor (EC 3.4.21.7) [Contains: Angiostatin].
GN PLG.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RN SEQUENCE FROM N.A.
RP MEDLINE=91184812; PubMed=2081600;
RX MEDLINE=91184812; PubMed=7525077;
RA Degen S.J., Bell S.M., Schaefer L.A., Elliott R.W.;
RT "Characterization of the cDNA coding for mouse plasminogen and
RT localization of the gene to mouse chromosome 17.";
RL Genomics 8:49-61(1990).
RN [2]
RN CHARACTERIZATION OF ANGIOSTATIN, AND PARTIAL SEQUENCE.
RX MEDLINE=95042728; PubMed=7525077;
RA O'Reilly M.S., Holmgren L., Shing Y., Chen C., Rosenthal R.A.,
RA Moses M., Lane W.S., Cao Y., Sage E.H., Folkman J.;
RT "Angiostatin: a novel angiogenesis inhibitor that mediates the
RT suppression of metastases by a Lewis lung carcinoma.";
RL Cell 79:315-328(1994).
CC -!- FUNCTION: PLASMIN DISSOLVES THE FIBRIN OF BLOOD CLOTS AND ACTS AS
CC A PROTEOLYTIC FACTOR IN A VARIETY OF OTHER PROCESSES INCLUDING
CC EMBRYONIC DEVELOPMENT, TISSUE REMODELING, TUMOR INVASION,
CC AND INFLAMMATION; IN OVULATION IT WEAKENS THE WALLS OF THE
CC GRAAFIAN FOLLICLE. IT ACTIVATES THE UROKINASE-TYPE PLASMINOGEN
CC ACTIVATOR, COLLAGENASES AND SEVERAL COMPLEMENT ZYMOGENS, SUCH
CC AS C1 AND C5. IT CLEAVES FIBRIN, FIBRONECTIN, THROMBOSPONDIN,
CC LAMININ AND VON WILLEBRAND FACTOR.
CC -!- FUNCTION: ANGIOSTATIN IS AN ANGIOGENESIS INHIBITOR THAT BLOCKS
CC NEOVASCULARIZATION AND GROWTH OF EXPERIMENTAL PRIMARY AND
CC METASTATIC TUMORS IN VIVO.
CC -!- CATALYTIC ACTIVITY: Preferential cleavage: Lys-|-Xaa > Arg-|-Xaa;
CC higher selectivity than trypsin. Converts fibrin into soluble
CC products.
CC -!- ENZYME REGULATION: CONVERTED INTO PLASMIN BY PLASMINOGEN
CC ACTIVATORS, BOTH PLASMINOGEN AND ITS ACTIVATOR BEING BOUND TO
CC FIBRIN. CANNOT BE ACTIVATED WITH STREPTOKINASE.
CC -!- MISCELLANEOUS: PLASMIN IS INACTIVATED BY ALPHA-2-ANTIPLASMIN
CC IMMEDIATELY AFTER DISSOCIATION FROM THE CLOT.
CC -!- MISCELLANEOUS: IN THE PRESENCE OF THE INHIBITOR, THE ACTIVATION
CC INVOLVES ONLY CLEAVAGE AFTER ARG-581, RESULTING IN 2 CHAINS HELD
CC TOGETHER BY 2 DISULFIDE BONDS. WITHOUT THE INHIBITOR, THE
CC ACTIVATION INVOLVES ALSO REMOVAL OF THE ACTIVATION PEPTIDE.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1. PLASMINOGEN SUBFAMILY.
CC -!- SIMILARITY: CONTAINS 5 KRINGLE DOMAINS.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@lsb-sib.ch).
-----
CC EMBL; J04766; AAA50168.1; -
CC PIR; A38514; A38514.
CC HSSP; P00747; 1PMK.
CC MEROPS; S01.233; -.
CC MGD; MGI:97620; Plg.
CC InterPro; IPR001314; Chymotrypsin.

```

```

DR InterPro: IPR000001; Kringle.
DR InterPro: IPR003014; PAN.
DR InterPro: IPR003609; Pan_app.
DR InterPro: IPR001254; Ser_protease_Try.
DR Pfam: PF00024; PAN; 1.
DR Pfam: PF00051; kringle; 5.
DR Pfam: PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 5.
DR SMART; SM00130; KR; 5.
DR SMART; SM00473; PAN_AP; 1.
DR SMART; SM00020; TRY_SPC; 1.
DR PROSITE; PS00021; KRINGLE_1; 4.
DR PROSITE; PS00070; KRINGLE_2; 5.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydrolase; Serine protease; Plasma; Glycoprotein; Fibrinolysis;
KW Tissue remodeling; Blood coagulation; Kringle; Zymogen; Repeat;
KW Signal.
FT SIGNAL 1 19
FT CHAIN 20 812 PLASMINOGEN.
FT CHAIN 20 581 PLASMIN HEAVY CHAIN A.
FT PEPTIDE 20 97 ACTIVATION PEPTIDE.
FT CHAIN 98 581 PLASMIN SHORT FORM OF CHAIN A.
FT CHAIN 98 7436 ANGIOSTATIN.
FT CHAIN 582 812 PLASMIN LIGHT CHAIN B.
FT DOMAIN 103 181 KRINGLE 1.
FT DOMAIN 184 262 KRINGLE 2.
FT DOMAIN 275 352 KRINGLE 3.
FT DOMAIN 377 454 KRINGLE 4.
FT DOMAIN 481 560 KRINGLE 5.
FT DOMAIN 582 812 SERINE PROTEASE.
FT ACT_SITE 624 624 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 667 667 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 762 762 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT DISULFID 49 73 BY SIMILARITY.
FT DISULFID 53 61 BY SIMILARITY.
FT DISULFID 103 181 BY SIMILARITY.
FT DISULFID 124 164 BY SIMILARITY.
FT DISULFID 152 176 BY SIMILARITY.
FT DISULFID 185 262 BY SIMILARITY.
FT DISULFID 188 316 BY SIMILARITY.
FT DISULFID 206 245 BY SIMILARITY.
FT DISULFID 234 257 BY SIMILARITY.
FT DISULFID 275 335 BY SIMILARITY.
FT DISULFID 296 335 BY SIMILARITY.
FT DISULFID 324 347 BY SIMILARITY.
FT DISULFID 377 454 BY SIMILARITY.
FT DISULFID 398 437 BY SIMILARITY.
FT DISULFID 426 449 BY SIMILARITY.
FT DISULFID 481 560 BY SIMILARITY.
FT DISULFID 502 543 BY SIMILARITY.
FT DISULFID 531 555 BY SIMILARITY.
FT DISULFID 568 687 INTERCHAIN (BY SIMILARITY).
FT DISULFID 578 586 INTERCHAIN (BY SIMILARITY).
FT DISULFID 609 625 BY SIMILARITY.
FT DISULFID 701 768 BY SIMILARITY.
FT DISULFID 731 747 BY SIMILARITY.
FT DISULFID 758 786 BY SIMILARITY.
SQ SEQUENCE 812 AA; 90846 MW; D34A74A4FC2256F8 CRC64;

```

Query Match 40.7%; Score 1059; DB 1; Length 812;  
 Best Local Similarity 43.3%; Pred.No. 1e-73;  
 Matches 191; Conservative 67; Mismatches 145; Indels 38; Gaps 12;  
 QY 30 TKKVNTA---DQCANRCTRNGKLPFTCKAFVDFKARQKQCLWPFNMSGGVKKFGEHED 86  
 DB 37 TKKQLAAGGVSDCLAKECE--TDFVCRSFQVHSKEQOCVMAENSKTSSIRF--RDVI 92  
 QY 87 LYENKDYIRNCIIGKGRSYKGVSTKSIKQCPQSSMIPHEHSLPSSYRGKDLQENVC 146

Db 93 LFEKRVYLSECKTGIGNGYRGTMTSGVACQKAGTTFPHVNPVSP5THPNEGLENYC 152

Qy 147 RNPGEEGPWCFTSNPEVRYEVDIPQCSVEVECMTCNGESYRGLMDHTSGKICQWRDH 206

Db 153 RNPNDDEGPCWYTTDPKDYDCNIPCEB-ECMYCSGEYKIGKISKITMSGLDQAWDS 211

Qy 207 QTPRHKFLPRLYDKGDFDDNYCRNPDGQPRWCYTLDPHTRWEYCAIKTCADNTMNDT 266

Db 212 QSPHAGYIPAKFTSKNLKMYCNPNPGEPRWCFTTDPTRWEYCDIPRCT-TPPP 267

Qy 267 VPLETTCIOGEGYRGYRTVNTIWNIPQCRWDQSYRGLMDHTPENFKCKDLRENYCRNP 326

Db 268 PPSPTYQCLKRGYRGTSTVTSYSGTKCQWSEQTPRHNRTPENFPCKNLENYCRNP 327

Qy 327 DGSSEPCWFTTDPNIRVGYCSQIPNCDMSHG-----QDCYRGNKNYM 369

Db 328 DGETAPWCYTTDSLRWEYC-EIPSCSSASPDQSDSSVPPEEQTPVVQECYQSDGQSYR 386

Qy 370 GNLQTSRGLTCSMWKDMEDLHRHIFWE---PDASKLNYCRNPDDDAHPWCYTCNP 426

Db 387 GTSSTTIITGKRCQSWAAMP--HRHSKTPENFPDAG-LEMNYCRNPDDG-KGPWCYTTDP 442

Qy 427 LIPWDCYPISRCEGDTTPTIV 447

Db 443 SVRWEYCNLKKC-SETGGSV 462

RESULT 8

PLMN\_ERIEU STANDARD; PRF: 810 AA.

AC Q29485;

DT 01-NOV-1997 (Rel. 35, Created)

DT 01-NOV-1997 (Rel. 35, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Plasminogen precursor (EC 3.4.21.7).

GN PLG.

OS Erinaceus europaeus (Western European hedgehog).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Insectivora; Erinaceidae; Erinaceinae; Erinaceus.

OX NCBI\_TaxID=9365;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Liver;

RX MEDLINE=96025778; PubMed=7592597;

RA Lawn R.M., Boommark N.W., Schwartz K., Lindahl G.E., Wade D.P.,

RA Byrne C.D., Fong K.J., Meer K., Patthy L.;

RT "The recurring evolution of lipoprotein(a)." Insights from cloning of

RT hedgehog apolipoprotein(a)."

RL J. Biol. Chem. 270:24004-24009(1995).

RN [2]

RP REVISIONS.

RA Lawn R.M.;

RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.

CC -1- FUNCTION: PLASMIN DISSOLVES THE FIBRIN OF BLOOD CLOTS AND ACTS AS

CC A PROTEOLYTIC FACTOR IN A VARIETY OF OTHER PROCESSES INCLUDING

CC EMBRYONIC DEVELOPMENT, TISSUE REMODELING, TUMOR INVASION,

CC AND INFLAMMATION; IN OVULATION IT WEAKENS THE WALLS OF THE

CC GRAAFIAN FOLLICLE. IT ACTIVATES THE UROKINASE-TYPE PLASMINOGEN

CC ACTIVATOR, COLLAGENASES AND SEVERAL COMPLEMENT ZYMOGENS. SUCH

CC AS C1 AND C5. IT CLEAVES FIBRIN, FIBRONECTIN, THROMBOSPONDIN,

CC LAMININ AND VON WILLEBRAND FACTOR.

CC -1- CATALYTIC ACTIVITY: Preferential cleavage: Lys-|-Xaa > Arg-|-Xaa;

CC higher selectivity than trypsin. Converts fibrin into soluble

CC products.

CC -1- ENZYME REGULATION: CONVERTED INTO PLASMIN BY PLASMINOGEN

CC ACTIVATORS, BOTH PLASMINOGEN AND ITS ACTIVATOR BEING BOUND TO

CC FIBRIN. CANNOT BE ACTIVATED WITH STREPTOKINASE.

CC -1- MISCELLANEOUS: PLASMIN IS INACTIVATED BY ALPHA-2-ANTIPLASMIN

CC IMMEDIATELY AFTER DISSOCIATION FROM THE CLOT.

CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1. PLASMINOGEN SUBFAMILY.

CC -1- SIMILARITY: CONTAINS 5 KRINGLE DOMAINS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

DR EMBL; U33171; AAC48717.1; -

DR HSP; P00747; LPMK.

DR MEROPS; S01.233; -

DR InterPro; IPR001314; Chymotrypsin.

DR InterPro; IPR000001; Kringle.

DR InterPro; IPR003014; PAN.

DR InterPro; IPR003609; Pan\_app.

DR InterPro; IPR001254; Ser\_protease\_Try.

DR Pfam; PF00024; PAN; 1.

DR Pfam; PF00051; kringle; 5.

DR Pfam; PF00089; trypsin; 1.

DR PRINTS; PR00722; CHYMOTRYPSIN.

DR PRINTS; PR00018; KRINGLE.

DR PRODOM; PD000395; Kringle; 5.

DR SMART; SM00130; KR; 5.

DR SMART; SM00473; PAN\_AP; 1.

DR SMART; SM00020; Tryp\_Spc; 1.

DR PROSITE; PS00021; KRINGLE\_1; 5.

DR PROSITE; PS00070; KRINGLE\_2; 5.

DR PROSITE; PS02040; TRYPsin\_DOM; 1.

DR PROSITE; PS00134; TRYPsin\_HIS; 1.

DR PROSITE; PS00135; TRYPsin\_SER; 1.

KW Hydrolase; Serine protease; Plasma; Glycoprotein; Fibrinolysis;

KW Tissue remodeling; Blood coagulation; Kringle; Zymogen; Repeat;

KW Signal.

FT CHAIN 1 19 BY SIMILARITY.

FT CHAIN 20 810 PLASMINOGEN.

FT CHAIN 20 582 PLASMIN HEAVY CHAIN A (BY SIMILARITY).

FT CHAIN 583 810 PLASMIN LIGHT CHAIN B (BY SIMILARITY).

FT DOMAIN 583 810 SERINE PROTEASE.

FT DOMAIN 103 181 KRINGLE 1.

FT DOMAIN 185 262 KRINGLE 2.

FT DOMAIN 275 352 KRINGLE 3.

FT DOMAIN 379 456 KRINGLE 4.

FT DOMAIN 482 561 KRINGLE 5.

FT ACT\_SITE 622 622 CHARGE RELAY SYSTEM.

FT ACT\_SITE 665 665 CHARGE RELAY SYSTEM.

FT ACT\_SITE 760 760 CHARGE RELAY SYSTEM.

FT CARBOHYD 339 339 N-LINKED (GLCNAC...) (POTENTIAL).

SQ SEQUENCE 810 AA; 90902 MW; 8E75780946017A16 CRC64;

Query Match 40.5%; Score 1053.5; DB 1; Length 810;

Best Local Similarity 44.3%; Pred. No. 2.7e-73;

Matches 194; Conservative 63; Mismatches 138; Indels 43; Gaps 13;

Qy 30 TKK---VNTADOCANRCTRNKGLPTCKAFVFDKARKOCLMFPFNSMSSGVKKFGEHFD 86

Db 37 TKKLSVSGTEBACVK--EKETSFICRSFYHSKQCVMAENSKSTPVLRM--RDVI 92

Qy 87 LYENKDYRNICIIIGKRSYKGTVTITKSGIKQCPMSMIPHSPLSPSYRKDLQENVC 146

Db 93 LFEKMYLSECKVNGKYYRGTVSKTKGLTCOKWSAETPHKPRESPDENEGLDQNYC 152

Qy 147 RNPGEEGPWCFTSNPEVRYEVDIPQCSVEVECMTCNGESYRGLMDHTSGKICQWRDH 206

Db 153 RNPNDDEGPCWYTTDPKDYDCNIPCEB-ECMYCSGEYKIGKISKITMSGLDQAWDS 211

Qy 207 QTPRHKFLPRLYDKGDFDDNYCRNPDGQPRWCYTLDPHTRWEYCAIKTCADNTMNDT 266

Db 212 QSPHAGYIPAKFTSKNLKMYCNPNPGEPRWCFTTDPTRWEYCDIPRCT-TPPP 267

Qy 267 VPLETTCIOGEGYRGYRTVNTIWNIPQCRWDQSYRGLMDHTPENFKCKDLRENYCRNP 326

Db 268 PPSPTYQCLKRGYRGTSTVTSYSGTKCQWSEQTPRHNRTPENFPCKNLENYCRNP 327

Qy 327 DGSSEPCWFTTDPNIRVGYCSQIPNCDMSHG-----QDCYRGNKNYM 369

Db 328 DGETAPWCYTTDSLRWEYC-EIPSCSSASPDQSDSSVPPEEQTPVVQECYQSDGQSYR 386

Qy 370 GNLQTSRGLTCSMWKDMEDLHRHIFWE---PDASKLNYCRNPDDDAHPWCYTCNP 426

Db 387 GTSSTTIITGKRCQSWAAMP--HRHSKTPENFPDAG-LEMNYCRNPDDG-KGPWCYTTDP 442

Qy 427 LIPWDCYPISRCEGDTTPTIV 447

Db 443 SVRWEYCNLKKC-SETGGSV 462



37 TKQLGANGSTEECAACEEEE--EFTCRSFQYHSKEQCVIAENRKSSIVFRM--RDVV 92

87 LYENKDYIRNCIIIGKRSYKGTVSIITKSGIKQCPWSSMIPHEHSFLPSSYRGKDLQENYC 146

93 LFEKKVYLSCKTGNGKNYRGWTSKTRTGTICQKWSSTSPHRPTFSPATHPSEGLEENYC 152

147 RNPGRBEGGPMWCTSNPEYRYEVCDDIPQCSVEVCWTCNGESYRGLMDHTESGKICQWDH 206

153 RNPNDQGGQPMWCTTDPERFDYCDIPEC-EDECMHCSGENYDGKTSKTMGLEQAWDS 211

207 QTPIRHKFLPERYPDKGFDNDNCRNPDGQRPWCYCTLDPHTRWEYCAIKTCADNTMNDTD 266

212 QSPHANGYIPSKFPNNLKNKYNCRNPDGGRPWCFTTDPNKKRWELCDIPRCT-----TP 265

267 VPLE--TTETCIQGGEGYRGYVNTIWNIGIPCQRWDQSYRPHEDMTPEPFCKDLRENYCR 324

266 PPSSGPTYQCLKGTGENYRGDVAVTVSGTHCHGWSAQTPTHNRTPENFPCKNLNDENYCR 325

325 NPDGSSPWCFTTDPNIRVGYQVQIINCDSH-----GQDCYRGNKGN 367

326 NPDGKAPWCYTTNSQVRWEYC-KIPSCSSPVSTPEFLDPTAPPELTPVVQCYCHGDGOS 384

368 YMGNLQSTRSGLTCSMW-----DKNMDLHRHFWEPPDASKNLNENYCRNPDDAHGP 419

385 YRGTSSTTTGKKCQSWSSMTPHWHKRTPENF-----PNAG-LPMNYCRNPDDAD-KGP 435

420 WCYTGNPLIPWDYCPISRCBG 440

436 WCFTTDPSPVRWEYCNLKKCSG 456

RESULT 10

ID	PLMN_HUMAN	STANDARD;	PRT;	810 AA.
AC	P00747;			
DT	21-JUL-1986 (Rel. 01, Created)			
DT	01-MAR-1989 (Rel. 10, Last sequence update)			
DT	15-JUN-2002 (Rel. 41, Last annotation update)			
DE	Plasminogen precursor (EC 3.4.21.7) [Contains: Angiostatatin].			
DE	PLG.			
GN	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RP	[1]			
RP	SEQUENCE FROM N.A.			
RA	MEDLINE=90202879; PubMed=2318848;			
RX	Petersen T.E., Martzen M.R., Ichinose A., Davie E.W.;			
RT	"Characterization of the gene for human plasminogen, a key proenzyme			
RT	in the fibrinolytic system.;"			
RL	J. Biol. Chem. 265:6104-6111(1990).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RP	MEDLINE=87162490; PubMed=3030813;			
RX	Forsgren M., Raden B., Israelsson M., Larsson K., Heden L.-O.;			
RT	"Molecular cloning and characterization of a full-length cDNA clone			
RT	for human plasminogen.;"			
RL	FEBS Lett. 213:254-260(1987).			
RN	[3]			
RP	SEQUENCE OF 20-810.			
RA	Sottrup-Jensen L., Petersen T.E., Magnusson S.;			
RX	Submitted (JUL-1977) to the PIR data bank.			
RN	[4]			
RP	SEQUENCE OF 292-810 FROM N.A.			
RX	MEDLINE=85023311; PubMed=6148961;			
RA	Wallinowski D.P., Sadler J.E., Davie E.W.;			
RT	"Characterization of a complementary deoxyribonucleic acid coding for			
RT	human and bovine plasminogen.;"			
RL	Biochemistry 23:4243-4250(1984).			
RN	[5]			
RP	SEQUENCE OF 20-100.			
RX	MEDLINE=75093329; PubMed=122932;			
RA	Wilan B., Wallen P.;			

"Structural relationship between 'glutamic acid' and 'lysine' forms of human plasminogen and their interaction with the NH2-terminal activation peptide as studied by affinity chromatography.";  
Eur. J. Biochem. 50:489-494(1975).  
[6]  
SEQUENCE OF 95-580; 581-626; 657-700 AND 732-810.  
Sottrup-Jensen L., Claeys H., Zajdel M., Petersen T.E., Magnusson S.;  
RA (In) Davidson J.F., Rowan R.M., Samama M.M., Desnoyers P.C. (eds.);  
RL Progress in chemical fibrinolysis and thrombolysis, pp.3:191-209,  
Raven Press, New York (1978).  
[7]  
SEQUENCE OF 483-604.  
MEDLINE=76043692; PubMed=126863;  
Raman B., Wallen P.;  
RA "Amino-acid sequence of the cyanogen-bromide fragment from human  
RL plasminogen that forms the linkage between the plasmin chains.";  
Eur. J. Biochem. 58:539-547(1975).  
[8]  
SEQUENCE OF 581-810.  
MEDLINE=77225245; PubMed=142009;  
Raman B.;  
RA "Primary structure of the B-chain of human plasmin.";  
Eur. J. Biochem. 76:129-137(1977).  
[9]  
ACTIVE SITE.  
MEDLINE=73149248; PubMed=4694729;  
RA Robbins K.C., Bernabe P., Arzadon L., Summaria L.;  
RL "The primary structure of human plasminogen. II. The histidine loop  
of human plasmin: light (B) chain active center histidine sequence.";  
J. Biol. Chem. 248:1631-1633(1973).  
[10]  
ACTIVE SITE.  
MEDLINE=69234739; PubMed=4240117;  
RA Groskopf W.R., Summaria L., Robbins K.C.;  
RL "Studies on the active center of human plasmin. Partial amino acid  
sequence of a peptide containing the active center serine residue.";  
J. Biol. Chem. 244:3590-3597(1969).  
[11]  
OMEGA-AMINOCARBOXYLIC ACID-BINDING SITES.  
MEDLINE=82213905; PubMed=6919539;  
RA Trexler M., Vail Z., Patthy L.;  
RL "Structure of the omega-aminocarboxylic acid-binding sites of human  
plasminogen. Arginine 70 and aspartic acid 56 are essential for  
binding of ligand by kringle 4.";  
J. Biol. Chem. 257:7401-7406(1982).  
[12]  
FIBRIN AND OMEGA-AMINOCARBOXYLIC ACID BINDING SITES.  
MEDLINE=85034794; PubMed=6094528;  
RA Vail Z., Patthy L.;  
RL "The fibrin-binding site of human plasminogen. Arginines 32 and 34  
are essential for fibrin affinity of the kringle 1 domain.";  
J. Biol. Chem. 259:13690-13694(1984).  
[13]  
PHOSPHORYLATION SITE SER-597.  
MEDLINE=97345939; PubMed=9201958;  
RA Wang H., Prorok M., Bretthauer R.K., Castellino F.J.;  
RL "Serine-578 is a major phosphorylation locus in human plasma  
plasminogen.";  
Biochemistry 36:8100-8106(1997).  
[14]  
CARBOHYDRATE-LINKAGE SITES.  
MEDLINE=88185329; PubMed=3356193;  
RA Marti T., Schaller J., Rickli E.E., Schmid K., Kamerling J.P.,  
RA Gerwig G.J., van Halbeek H., Vliegthart J.F.;  
RL "The N- and O-linked carbohydrate chains of human, bovine and porcine  
plasminogen. Species specificity in relation to sialylation and  
fucosylation patterns.";  
Eur. J. Biochem. 173:57-63(1988).  
[15]  
CARBOHYDRATE-LINKAGE SITE SER-268.  
MEDLINE=97207306; PubMed=9054441;  
RA Pirie-Shepherd S.R., Stevens R.D., Andon N.L., Enghild J.J.,  
RA Pizzo S.V.;

RT "Evidence for a novel O-linked sialylated trisaccharide on Ser-248 of human plasminogen 2.";  
RL J. Biol. Chem. 272:7408-7411(1997).  
RN [16]  
RP CHARACTERIZATION OF ANGIOTATIN, AND PARTIAL SEQUENCE.  
RX MEDLINE-95042728; PubMed-7525077;  
RA O'Reilly M.S., Holmgren L., Shing Y., Chen C., Rosenthal R.A.,  
RT Moses M., Lane W.S., Cao Y., Sage E.H., Folkman J.;  
RT "Angiostatin: a novel angiogenesis inhibitor that mediates the  
RT suppression of metastases by a Lewis lung carcinoma.";  
RL Cell 79:315-328(1994).  
RN [17]  
RP CHARACTERIZATION OF ANGIOTATIN.  
RX MEDLINE-97238710; PubMed-9102221;  
RA Sim B.K., O'Reilly M.S., Liang H., Fortier A.H., He W., Madsen J.W.,  
RL Laptevich R., Nacy C.A.;  
RT "A recombinant human angiotensin protein inhibits experimental primary  
RT and metastatic cancer.";  
RL Cancer Res. 57:1329-1334(1997).  
RN [18]  
RP X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS) OF 374-461.  
RX MEDLINE-92031502; PubMed-1657148;  
RA Mulichak A.M., Tulinsky A., Ravichandran K.G.;  
RT "Crystal and molecular structure of human plasminogen kringles 4  
RT refined at 1.9-A resolution.";  
RL Biochemistry 30:10576-10588(1991).  
RN [19]  
RP X-RAY CRYSTALLOGRAPHY (2.25 ANGSTROMS) OF 374-461.  
RX MEDLINE-92031503; PubMed-1657149;  
RA Wu T.-P., Padmanabhan K., Tulinsky A., Mulichak A.M.;  
RT "The refined structure of the epsilon-aminocaproic acid complex of  
RT human plasminogen kringles 4.";  
RL Biochemistry 30:10589-10594(1991).  
RN [20]  
RP X-RAY CRYSTALLOGRAPHY (1.67 ANGSTROMS) OF 376-454.  
RX MEDLINE-96180681; PubMed-8611560;  
RA Mathews I.I., Vanderhoff-Hanover P., Castellino F.J., Tulinsky A.;  
RT "Crystal structures of the recombinant kringles 1 domain of human  
RT plasminogen in complexes with the ligands epsilon-aminocaproic acid  
RT and trans-4-(aminomethyl)cyclohexane-1-carboxylic Acid.";  
RL Biochemistry 35:2567-2576(1996).  
RN [21]  
RP X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS) OF 102-181.  
RX MEDLINE-96180681; PubMed-8611560;  
RA Mathews I.I., Vanderhoff-Hanover P., Castellino F.J., Tulinsky A.;  
RT "Crystal structures of the recombinant kringles 1 domain of human  
RT plasminogen in complexes with the ligands epsilon-aminocaproic acid  
RT and trans-4-(aminomethyl)cyclohexane-1-carboxylic Acid.";  
RL Biochemistry 35:2567-2576(1996).  
RN [22]  
RP X-RAY CRYSTALLOGRAPHY (1.66 ANGSTROMS) OF 480-563.  
RX MEDLINE-98198034; PubMed-9521645;  
RA Chang Y., Mochalkin I., McCance S.G., Cheng B., Tulinsky A.,  
RT Castellino F.J.;  
RT "Structure and ligand binding determinants of the recombinant kringles  
RT 5 domain of human plasminogen.";  
RL Biochemistry 37:3258-3271(1998).  
RN [23]  
RP STRUCTURE BY NMR OF 96-184.  
RX MEDLINE-94237157; PubMed-8181475;  
RA Rejzante M.R., Llinas M.;  
RT "1H-NMR assignments and secondary structure of human plasminogen  
RT kringles 1.";  
RL Eur. J. Biochem. 221:927-937(1994).  
RN [24]  
RP STRUCTURE BY NMR OF 96-184.  
RX MEDLINE-94237158; PubMed-8181476;  
RA Rejzante M.R., Llinas M.;  
RT "Solution structure of the epsilon-aminohexanoic acid complex of  
RT human plasminogen kringles 1.";  
RL Eur. J. Biochem. 221:939-949(1994).  
RN [25]  
RP STRUCTURE BY NMR OF 183-354.  
RX MEDLINE-96194156; PubMed-8652577;  
RA Soehndel S., Hu C.-K., Marti D., Affolter M., Schaller J., Llinas M.,

RA Rickli E.E.;  
RT "Recombinant gene expression and 1H NMR characteristics of the  
RT kringles (2 + 3) supermodule: spectroscopic/functional individuality  
RT of plasminogen kringles domains.";  
RL Biochemistry 35:2357-2364(1996).  
RN [26]  
RP STRUCTURE BY NMR OF 374-461.  
RX MEDLINE-90219023; PubMed-2157850;  
RA Atkinson R.A., Williams R.J.P.;  
RT "Solution structure of the kringles 4 domain from human plasminogen by  
RT 1H nuclear magnetic resonance spectroscopy and distance geometry.";  
RL J. Mol. Biol. 212:541-552(1990).  
RN [27]  
RP VARIANTS PHE-374 AND THR-620.  
Query Match 40.0%; Score 1042; DB 1; Length 810;  
Best Local Similarity 43.8%; Pred. No 2e-72;  
Matches 193; Conservative 68; Mismatches 140; Indels 42; Gaps 13;  
QY 30 TKK---VNTADQANCRNRKGLPFTCKAFVFDKARKQCLWFFNFNSMSSGVKKFGEHFD 86  
DB 37 TKKQLGAGSIECAKCEDE--EFTCRAFOYHSKEQCVMAENRRKSSIIIRM--RDVV 92  
QY 87 LYENKDIYRNCIIGKRSYKGTVSITKSGIKCQWPSSMIPEHSEFLPSSYRGDLQENYC 146  
DB 93 LFEKKVYLSECKTGNGKNYRGTMSKTNGITCQKWSSTSPHRPFPSPATHPSGLEENYC 152  
QY 147 RNPREGGEGPWCFTSNPEVRYEVCDDIPQCSEVECMTCGSEYRGLMDHTESGKICQWDH 206  
DB 153 RNPNDPQGPWCYTDPKRYDYCDILECEE--ECMHCSGENYDGIKSTMSGLECQAWDS 211  
QY 207 QTPHRHKLPERYDPKGFDDNYCRNPDGQPPWCYTILDPHTRWEYCAIKTCADNTMNDT 266  
DB 212 QSPHAGVYIPSKFNNKLNKKNYCRNPDRELPRWCFITDPNKNWELCDIPRCT-----TP 265  
QY 267 VPLE--TTECIQOGGEGYRGVTIWNIGICQWRDQVPEHDMTPNFCKDLRENYCR 324  
DB 266 PPSGPTTQCLKRGTEGNGVAVTVSGHTCQHSQAQTHNTPTNFPCKNLDENYCR 325  
QY 325 NPDGSESPWCFTTDFNIRVGYCSQIPNCDMSH-----GDCYRGNGKN 367  
DB 326 NPDGKRAPWCHTNSQVWEYC-KIPSCDSPVSTEQLAPPELTPVVDQCYHGDQS 384  
QY 368 YGNLSQTRSLGTSMDKMDLHRHFWF--PDASKLNEYCRNPDADDHGPWCYT 424  
DB 385 YGTSSTTTTGKCKOSWSSMTP--HRHOKTPENYPNAG-LTMNYCRNPDAD-KGPWCFTT 440  
QY 425 NPLIPWDYCPISRCGDTPTIV 447  
DB 441 DPSVWEYCNLKKCSG-TEASVW 462  
RESULT 11  
ID PLMN BOVIN STANDARD; PRT; 812 AA.  
AC P06868; Q28162;  
DT 01-JAN-1988 (Rel. 06, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Plasminogen precursor (EC 3.4.21.7).  
GN PLG.  
OS Bos taurus (Bovine).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
OC Bovidae; Bovinae; Bos.  
OX NCBI\_TaxID=9913;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Liver;  
RA Berglund L., Andersen M.D., Petersen T.E.;  
RT "Cloning and characterization of the bovine plasminogen cDNA.";  
RL Int. Dairy J. 5:593-603(1995).  
RN [2]



RP SEQUENCE OF 27-812, AND CARBOHYDRATE-LINKAGE SITES.  
RX MEDLINE=85203906; PubMed=3846532;  
RA Schaller J., Moser P.W., Danneberg-Muller G.A.K., Rosselet S.J.,  
RT Kamfer U., Rickli E.E.;  
RT "Complete amino acid sequence of bovine plasminogen. Comparison with  
RL human plasminogen."; Eur. J. Biochem. 149:267-278(1985).  
RN [3]  
RP SEQUENCE OF 706-812 FROM N.A.  
RX MEDLINE=85023311; PubMed=6148961;  
RA Malinowski D.P., Sadler J.E., Davie E.W.;  
RT "Characterization of a complementary deoxyribonucleic acid coding for  
RL human and bovine plasminogen."; Biochemistry 23:4243-4250(1984).  
RN [4]  
RP CARBOHYDRATE-LINKAGE SITES.  
RX MEDLINE=88185329; PubMed=3356193;  
RA Marti T., Schaller J., Rickli E.E., Schmid K., Kamerling J.P.,  
RA Gerwig G.J., van Halbeek H., Vliegenthart J.F.;  
RT "The N- and O-linked carbohydrate chains of human, bovine and porcine  
RT plasminogen. Species specificity in relation to sialylation and  
RL fucosylation patterns."; Eur. J. Biochem. 173:57-63(1988).  
CC -!- FUNCTION: PLASMIN DISSOLVES THE FIBRIN OF BLOOD CLOTS AND ACTS AS  
CC A PROTEOLYTIC FACTOR IN A VARIETY OF OTHER PROCESSES INCLUDING  
CC EMBRYONIC DEVELOPMENT, TISSUE REMODELING, TUMOR INVASION,  
CC AND INFLAMMATION; IN OVULATION IT WEAKENS THE WALLS OF THE  
CC GRAAFIAN FOLLICLE. IT ACTIVATES THE UROKINASE-TYPE PLASMINOGEN  
CC ACTIVATOR, COLLAGENASES AND SEVERAL COMPLEMENT ZYMOGENS. SUCH  
CC AS C1 AND C5. IT CLEAVES FIBRIN, FIBRONECTIN, THROMBOSPONDIN,  
CC LAMININ AND VON WILLEBRAND FACTOR.  
CC -!- CATALYTIC ACTIVITY: Preferential cleavage: Lys-|-Xaa > Arg-|-Xaa;  
CC higher selectivity than trypsin. Converts fibrin into soluble  
CC products.  
CC -!- ENZYME REGULATION: CONVERTED INTO PLASMIN BY PLASMINOGEN  
CC ACTIVATORS. BOTH PLASMINOGEN AND ITS ACTIVATOR BEING BOUND TO  
CC FIBRIN. CANNOT BE ACTIVATED WITH STREPTOKINASE.  
CC -!- PM- N-LINKED GLYCANS CONTAIN N-ACETYLGLUCOSAMINE AND SIALIC ACID.  
CC O-LINKED GLYCANS CONSIST OF GAL-GALNAc DISACCHARIDE WITH IS  
CC MODIFIED WITH UP TO 2 SIALIC ACID RESIDUES (MICROHETEROGENEITY).  
CC -!- MISCELLANEOUS: PLASMIN IS INACTIVATED BY ALPHA-2-ANTIPLASMIN  
CC IMMEDIATELY AFTER DISSOCIATION FROM THE CLOT.  
CC -!- SIMILARITY: BELONGS TO PERTIDASE FAMILY S1. PLASMINOGEN SUBFAMILY.  
CC -!- SIMILARITY: CONTAINS 5 KRINGLE DOMAINS.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL: X79402; CAA55939.1; -;  
DR EMBL: K02935; AAA30714.1; -;  
DR PIR: A25835; PLBO.  
DR HSP: P00747; 2PK4.  
DR MEROPS: S01.233; -;  
DR GlycoSuiteDB: P06868; -;  
DR InterPro: IPR001314; Chymotrypsin.  
DR InterPro: IPR000001; Kringle.  
DR InterPro: IPR003014; PAN.  
DR InterPro: IPR003609; Pan\_app.  
DR InterPro: IPR001254; Ser\_protease\_Try.  
DR Pfam: PF00024; PAN; 1.  
DR Pfam: PF00051; kringle; 5.  
DR Pfam: PF00089; trypsin; 1.  
DR PRINTS: PR00722; CHYMOTRYPSIN.  
DR PRINTS: PR00018; KRINGLE.  
DR ProDom: PD000395; kringle; 5.  
DR SMART: SM00130; KR; 5.  
DR SMART: SM00473; PAN.AP; 1.  
DR SMART: SM00020; Tryp\_Spc; 1.

DR PROSITE; PS00021; KRINGLE\_1; 5.  
DR PROSITE; PS00070; KRINGLE\_2; 5.  
DR PROSITE; PS50240; TRYPSIN\_DOM; 1.  
DR PROSITE; PS00134; TRYPSIN\_HIS; 1.  
DR PROSITE; PS00135; TRYPSIN\_SER; 1.  
KW Hydrolase; Serine protease; Plasmas; Glycoprotein; Fibrinolysis;  
KW Tissue remodeling; Blood coagulation; Kringle; Zymogen; Repeat;  
KW Signal.  
FT SIGNAL 1 26  
FT CHAIN 27 812 PLASMINOGEN  
FT CHAIN 27 583 PLASMIN HEAVY CHAIN A.  
FT CHAIN 584 812 PLASMIN LIGHT CHAIN B.  
FT DOMAIN 110 188 KRINGLE 1.  
FT DOMAIN 192 269 KRINGLE 2.  
FT DOMAIN 282 359 KRINGLE 3.  
FT DOMAIN 384 461 KRINGLE 4.  
FT DOMAIN 485 564 KRINGLE 5.  
FT DOMAIN 584 812 SERINE PROTEASE.  
FT CARBOHYD 315 315 N-LINKED (GLCNAC. .).  
FT CARBOHYD 365 365 /FTID-CAR-000014.  
FT CARBOHYD 624 624 O-LINKED (GALNAc. .).  
FT ACT\_SITE 624 624 /FTID-CAR-000015.  
FT ACT\_SITE 667 667 CHARGE RELAY SYSTEM.  
FT ACT\_SITE 762 762 CHARGE RELAY SYSTEM.  
FT CONFLICT 335 335 N -> D (IN REF. 2).  
FT CONFLICT 516 516 Q -> H (IN REF. 2).  
FT CONFLICT 555 555 P -> L (IN REF. 2).  
FT CONFLICT 744 744 T -> R (IN REF. 3).  
SQ SEQUENCE 812 AA; 91216 MW; 38A6AA691E220946 CRC64;  
  
Query Match 39.9%; Score 1038; DB 1; Length 812;  
Best Local Similarity 42.4%; Pred. No. 4.1e-72;  
Matches 185; Conservative 72; Mismatches 141; Indels 38; Gaps 11;  
  
QY 25 ALKIKTKVNTADQCACTRNKGLPFTCKAFVDFKARKQCLWFFPNSSSGVKKFEGHE 84  
DB 42 SLSRNLAGRSVEDCAAKC--EEETDFVCRAFOYHSKEQCVVMAENSKNTVPFRM--RD 97  
QY 85 FLYENKDYIRNCITIGKRSYKGTVSITKSGIKCOPWSSMIPHSFLPSSYRGKDLQEN 144  
DB 98 VLYEKRIYILLECKTNGOYRGTTAETKSGVTCOKWSATSPHVKFSPEKPLAGLEN 157  
QY 145 YCRNPRGEGGFWCTSNPEVRYEYCDIPQCSEVCEMTCNGESYRGLMDHTSGKICQRW 204  
DB 158 YCRNPNDENGPMCYTDPDKRYDYCDIPEC-EDKCMHCSGENYEGKIATKMSGRDQAW 216  
QY 205 DHQTPRHKFLPERYPDKGFDNDYCRNPDGQPRPCYTLDPHTRWYCAIKTCADNTMND 264  
DB 217 DSQSPHAHGYIPSKFPKNLKNMNYCRNPDGEPRPWCFTTDPQKRWEFCIDPCT----- 270  
QY 265 TDVPLE--TTECIGOGEGYRGTVNTIWNIPQCRWDQYPIEHQDWTPENFCKDKLRENY 322  
DB 271 TFPSSGPKYQCLKGTGKNGYGVAVTESGHTCQRWSEOTPHKHNPENPFCKNLEENY 330  
QY 323 CRNPDGSESPWCTTDPNIRVGYCSOIPNCDMSH-----GQCYRNG 365  
DB 331 CRNPGKAPWCYTINSEVREYCT-IPSCSSPLSTERMDVPVPEQTPVQDCYHGNG 389  
QY 366 KNYMGNLSQTRSGLTCSMDKNMEDLHRHFWB---PDASKLNENYCRNPDADAHCPWCY 422  
DB 390 QSYRTGTSSTIITGRKQCSWSSMTP--HRLKTPENYVNAAG-LTMNYCRNPAD-KSPWCY 445  
QY 423 TGNPLIPWDYCPISRC 438  
DB 446 TTDPVRWEFCNLKCC 461  
  
RESULT 12  
APOA\_HUMAN  
ID APOA\_HUMAN STANDARD; PRT; 4548 AA.  
AC P08519;  
DT 01-AUG-1988 (Rel. 08, Created)

DT 01-AUG-1988 (Rel. 08, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Apolipoprotein(a) precursor (BC 3.4.21.-) (Apo(a)) (Up(a)).  
GN LPA.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=88039109; PubMed=3670400;  
RA McLean J.W., Tomlison J.E., Kuang W.-J., Eaton D.L., Chen E.Y.,  
RA Fless G.M., Scanu A.M., Lawn R.M.;  
RT "cDNA sequence of human apolipoprotein(a) is homologous to  
RT plasminogen.";  
RL Nature 330:132-137(1987).  
RN [2]  
RP SERINE PROTEASE ACTIVITY.  
RX MEDLINE=90076123; PubMed=2531657;  
RA Salonen E.-M., Jauhainen M., Zardi L., Vaheri A., Ehnholm C.;  
RT "Lipoprotein(a) binds to fibronectin and has serine proteinase  
RT activity capable of cleaving it.";  
RL EMBO J. 8:4035-4040(1989).  
RN [3]  
RP REVIEW.  
RX MEDLINE=90049223; PubMed=2530631;  
RA Utermann G.;  
RT "The mysteries of lipoprotein(a).";  
RL Science 246:904-910(1989).  
RN [4]  
RP CHARACTERIZATION OF THE N- AND O-LINKED GLYCANS.  
RX MEDLINE=21303595; PubMed=11294842;  
RA Garner B., Merry A.H., Royle L., Harvey D.J., Rudd P.M., Thillet J.;  
RT "Structural elucidation of the N- and O-glycans of human  
RT apolipoprotein(a): role of o-glycans in conferring protease  
RT resistance.";  
RL J. Biol. Chem. 276:22200-22208(2001).  
RN [5]  
RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF 4121-4208.  
RX MEDLINE=96217891; PubMed=8642595;  
RA Mikol V., Lograsso P.V., Boettcher B.R.;  
RT "Crystal structures of apolipoprotein(a) kringle IV37 free and  
RT complexed with 6-aminohexanoic acid and with p-aminomethylbenzoic  
RT acid: existence of novel and expected binding modes.";  
RL J. Mol. Biol. 256:751-761(1996).  
RN [6]  
RP VARIANT ARG-4193.  
RX MEDLINE=95002201; PubMed=7918682;  
RA Scanu A.M., Pfaffinger D., Lee J.C., Hinman J.;  
RT "A single point mutation (Trp72-->Arg) in human apo(a) kringle 4-37  
RT associated with a lysine binding defect in Lp(a).";  
RL Biochim. Biophys. Acta 1227:41-45(1994).  
CC -!- FUNCTION: Apo(a) is the main constituent of lipoprotein(a)  
CC (Lp(a)). It has serine proteinase activity and is able of  
CC autolysis, inhibits tissue-type plasminogen activator 1.  
CC Lp(a) may be a ligand for megalin/Sp 330.  
CC -!- SUBUNIT: Disulfide-linked to apo-B100. Binds to fibronectin and  
CC decorin.  
CC -!- PTM: N- and O-glycosylated. The N-glycans are complex biantennary  
CC structures present in either a mono- or disialylated state. The  
CC O-glycans are mostly (80%) represented by the monosialylated core  
CC type I structure, NeuNAc2-3GalNAc-6S, with smaller  
CC amounts of disialylated and non-sialylated O-glycans also  
CC detected.  
CC -!- DISEASE: Elevated plasma concentrations of apo(a) and its  
CC naturally occurring proteolytic fragments are correlated with  
CC atherosclerosis. Homology with plasminogen kringle IV and V is  
CC thought to underlie the atherogenicity of the protein, because the  
CC fragments are competing with plasminogen for fibrin(ogen) binding.  
CC -!- MISCELLANEOUS: Apo(a) is known to be proteolytically cleaved,  
CC leading to the formation of the so called mini-Lp(a). Apo(a)  
CC fragments accumulate in atherosclerotic lesions, where they may  
CC promote thrombogenesis. O-glycosylation may limit the extent of

CC proteolytic fragmentation.  
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1. PLASMINOGEN SUBFAMILY.  
CC -!- SIMILARITY: CONTAINS 38 KRINGLE DOMAINS.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL: X06290; CAA29618.1; -.  
CC PIR: S00657; S00657.  
CC HSP: P00747; IPMK.  
CC MEROPS: S01.226; -.  
CC Genew: HGNC:6667; LPA.  
CC MIM: 152200; -.  
CC InterPro: IPR001314; Chymotrypsin.  
CC InterPro: IPR000001; Kringle.  
CC InterPro: IPR001254; Ser-protease\_Try.  
CC Pfam: PF00051; kringle; 38.  
CC Pfam: PF00089; trypsin; 1.  
CC PRINTS: PR00722; CHYMOTRYPSIN.  
CC PRINTS: PR00018; KRINGLE.  
CC ProDom: PD000395; Kringle; 38.  
CC SMART: SM00130; KR; 38.  
CC SMART: SM00020; Tryp-Spc; 1.  
CC PROSITE: PS00021; KRINGLE\_1; 38.  
CC PROSITE: PS50070; KRINGLE\_2; 38.  
CC PROSITE: PS50240; TRYPSIN\_DOM; 1.  
CC PROSITE: PS00134; TRYPSIN\_HIS; 1.  
CC PROSITE: PS00135; TRYPSIN\_SER; 1.  
CC Hydrolase: Serine protease; Lipid transport; Plasma; Glycoprotein;  
CC Kringle; Repeat; Atherosclerosis; Signal; Polymorphism.  
CC SIGNAL 1 19  
CC CHAIN 20 4548 APOLIPOPROTEIN(A).  
CC DOMAIN 20 130 KRINGLE TYPE IV, 1.  
CC DOMAIN 131 244 KRINGLE TYPE IV, 2.  
CC DOMAIN 245 358 KRINGLE TYPE IV, 3.  
CC DOMAIN 359 472 KRINGLE TYPE IV, 4.  
CC DOMAIN 473 586 KRINGLE TYPE IV, 5.  
CC DOMAIN 587 700 KRINGLE TYPE IV, 6.  
CC DOMAIN 701 814 KRINGLE TYPE IV, 7.  
CC DOMAIN 815 928 KRINGLE TYPE IV, 8.  
CC DOMAIN 929 1042 KRINGLE TYPE IV, 9.  
CC DOMAIN 1043 1156 KRINGLE TYPE IV, 10.  
CC DOMAIN 1157 1270 KRINGLE TYPE IV, 11.  
CC DOMAIN 1271 1384 KRINGLE TYPE IV, 12.  
CC DOMAIN 1385 1498 KRINGLE TYPE IV, 13.  
CC DOMAIN 1499 1612 KRINGLE TYPE IV, 14.  
CC DOMAIN 1613 1726 KRINGLE TYPE IV, 15.  
CC DOMAIN 1727 1840 KRINGLE TYPE IV, 16.  
CC DOMAIN 1841 1954 KRINGLE TYPE IV, 17.  
CC DOMAIN 1955 2068 KRINGLE TYPE IV, 18.  
CC DOMAIN 2069 2182 KRINGLE TYPE IV, 19.  
CC DOMAIN 2183 2296 KRINGLE TYPE IV, 20.  
CC DOMAIN 2297 2410 KRINGLE TYPE IV, 21.  
CC DOMAIN 2411 2524 KRINGLE TYPE IV, 22.  
CC DOMAIN 2525 2638 KRINGLE TYPE IV, 23.  
CC DOMAIN 2639 2752 KRINGLE TYPE IV, 24.  
CC DOMAIN 2753 2866 KRINGLE TYPE IV, 25.  
CC DOMAIN 2867 2980 KRINGLE TYPE IV, 26.  
CC DOMAIN 2981 3094 KRINGLE TYPE IV, 27.  
CC DOMAIN 3095 3208 KRINGLE TYPE IV, 28.  
CC DOMAIN 3209 3322 KRINGLE TYPE IV, 29.  
CC DOMAIN 3323 3436 KRINGLE TYPE IV, 30.  
CC DOMAIN 3437 3550 KRINGLE TYPE IV, 31.  
CC DOMAIN 3551 3664 KRINGLE TYPE IV, 32.  
CC DOMAIN 3665 3770 KRINGLE TYPE IV, 33.  
CC DOMAIN 3771 3884 KRINGLE TYPE IV, 34.  
CC DOMAIN 3885 3998 KRINGLE TYPE IV, 35.  
CC DOMAIN 3999 4112 KRINGLE TYPE IV, 36.

```
FT DOMAIN 4113 4226 KRINGLE TYPE IV, 37.
FT DOMAIN 4227 4327 KRINGLE TYPE V.
FT DOMAIN 4328 4548 SERINE PROTEASE.
FT ACT_SITE 4369 4369 CHARGE RELAY SYSTEM.
FT ACT_SITE 4412 4412 CHARGE RELAY SYSTEM.
FT ACT_SITE 4498 4498 CHARGE RELAY SYSTEM.
FT VARIANT 4193 4193 W -> R (LOSS OF LYSINE-SEPHAROSE BINDING).
FT SEQUENCE 4548 AA; 501313 MW; 96921BE96A465C5F CRC64;
Query Match 33.88; Score 879; DB 1; Length 4548;
Best Local Similarity 39.28; Pred. No. 4e-59;
Matches 166; Conservative 43; Mismatches 138; Indels 76; Gaps 7;
QY 89 ENKDYIRNCIIIGKRSYKGTVSITKGIKOPWSSMIPHEHSFLPSSYRGKDLQENYCRN 148
DB 3888 ENSTGVQDCYRGDQSGYRGTLSTITGRTQSWSSMTPHWHRRILYYPNAGLTRYCRN 3947
QY 149 PRGEGGPGWCFSTSNPEVRYEYCDIPOGSEV----- 179
DB 3948 PDABEIRPCWYTMDSYRWEYCNLTGCPVTSSVLTTPVAVPSTEAPSEQAPEKSPV 4006
QY 180 ---CWTNGESYRGLMDHTESGKTCQRWDHOTPHRHFLPERYPDKGFDNDCYCRNPQGP 236
DB 4007 VQDCYHGDGRSYRGSISSTVTGRTQSWSSMTPHWHQRTPENYPNAGLTENYCRNPDSGK 4066
QY 237 RPWCYTLDPHTRWEYCAIKTCADNTM---NDTDVPLETTE-----CQI 276
DB 4067 QPWCYTLDPCVRYEYCNLTGSETSEVLTGTPVPVPSMEAHSEAAPEQTGTPVVRQCYH 4126
QY 277 GQSGYRGTVNTIWNIGPCQRWDQSYRPHEDMTPEFKCKDLRENYCRNPDSGSPWCFT 336
DB 4127 GNGQSYRGTFSTVTGRTQSWSSMTPHRHQRTPENYPNDGLTWNCRNPDDTGPWCFT 4186
QY 337 TDPNIRVGYCS-----QIPNCDMHSQDCYRGNGKKNYGNLSQTRSG 378
DB 4187 MDPSIRWEYCNLTGSDTEGTVVAPPTVQVPSLGPPSEODCMFGNGKGYRGKATTVTG 4246
QY 379 LTCSMDKMNEDLHRLFWPEADSK---LNENYCRNPDDDAHGPWCYTGPNLIPWDCPI 435
DB 4247 TPCQEW---AAQEPHRHSTFTPGTNKAGLEKNYCRNPDDGNGPWCYTMNPKLFEDYCDI 4304
QY 436 SRC 438
DB 4305 PLC 4307
RESULT 13
APOA_MACMU STANDARD; PRT; 1420 AA.
AC P14417;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE APOlipoprotein(a) (EC 3.4.21.-) (Apo(a)) (Lp(a)) (Fragment).
GN LFA.
OS Macaca mulatta (Rhesus macaque).
OC Eukaryota; Metazoa; Chordata; Graniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecoidea; Macaca.
OX NCBI_TaxID=9544;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=89174660; PubMed=2925643;
RT Tomlinson J.E., McLean J.W., Lavin R.M.;
RT "Rhesus monkey apolipoprotein(a). Sequence, evolution, and sites of
RT synthesis."
RL J. Biol. Chem. 264:5957-5965(1989).
CC -1- FUNCTION: Apo(a) is the main constituent of lipoprotein(a)
CC (Lp(a)). It has serine proteinase activity and is able of
CC autoproteolysis. Inhibits tissue-type plasminogen activator 1.
CC Lp(a) may be a ligand for megalin/Gp 330.
```



RA Park C.H., Tulinsky A.;  
 RT "Three-dimensional structure of the kringle sequence: structure of  
 RT prothrombin fragment 1.";  
 RL Biochemistry 25:3977-3982(1986).  
 RN [5]  
 RP X-RAY CRYSTALLOGRAPHY (2.25 ANGSTROMS) OF ACTIVATION PEPTIDE 1.  
 RX MEDLINE-91311686; PubMed-1856863;  
 RA Seshadri T.-P., Tulinsky A., Skrzypczak-Jankun E., Park C.H.;  
 RT "Structure of bovine prothrombin fragment 1 refined at 2.25-A  
 RT resolution.";  
 RL J. Mol. Biol. 220:481-494(1991).  
 RN [6]  
 RP X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS) OF ACTIVATION PEPTIDE 1.  
 RX MEDLINE-92190185; PubMed-1547238;  
 RA Soriano-Garcia M., Padmanabhan K., de Vos A.M., Tulinsky A.;  
 RT "The Ca2+ ion and membrane binding structure of the Gla domain of Ca-  
 RT prothrombin fragment 1.";  
 RL Biochemistry 31:2554-2566(1992).  
 RN [7]  
 RP X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS).  
 RX MEDLINE-92218459; PubMed-1560020;  
 RA Martin P.D., Robertson W., Turk D., Huber R., Bode W., Edwards B.F.P.;  
 RT "The structure of residues 7-16 of the A alpha-chain of human  
 RT fibrinogen bound to bovine thrombin at 2.3-A resolution.";  
 RL J. Biol. Chem. 267:7911-7920(1992).  
 RN [8]  
 RP X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS).  
 RX MEDLINE-92389319; PubMed-1518046;  
 RA Brandstetter H., Turk D., Hoeffken H.W., Grosse D., Stuerzebecher J.,  
 RA Martin P.D., Edwards B.F.P., Bode W.;  
 RT "Refined 2.3 A x-ray crystal structure of bovine thrombin complexes  
 RT formed with the benzamide and arginine-based thrombin inhibitors  
 RT NAPAP, 4-TAPAP and MQPA. A starting point for improving  
 RT antithrombotics.";  
 RL J. Mol. Biol. 226:1085-1089(1992).  
 RN [9]  
 RP X-RAY CRYSTALLOGRAPHY (3.1 ANGSTROMS) OF COMPLEX WITH ORNITHODORIN.  
 RX MEDLINE-97102783; PubMed-8947023;  
 RA van de Loch A., Stubbs M.T., Bode W., Friedrich T., Bollschweiler C.,  
 RA Hoeffken W., Huber R.;  
 RT "The ornithodorin-thrombin crystal structure, a key to the TAP  
 RT enigma?";  
 RL EMBO J. 15:6011-6017(1996).  
 RN [10]  
 RP X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS) OF COMPLEX WITH TRIABIN.  
 RX MEDLINE-98004486; PubMed-9342325;  
 RA Fuentes-Prior P., Noeske-Jungblut C., Donner P., Schleuning W.D.,  
 RA Huber R., Bode W.;  
 RT "Structure of the thrombin complex with triabin, a lipocalin-like  
 RT exosite-binding inhibitor derived from a triatomine bug.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 94:11845-11850(1997).  
 RN [11]  
 RP GENE STRUCTURE.  
 RX MEDLINE-86077733; PubMed-3000440;  
 RA Irwin D.M., Ahern K.G., Pearson G.D., McGillivray R.T.A.;  
 RT "Characterization of the bovine prothrombin gene.";  
 RL Biochemistry 24:6854-6861(1985).  
 CC -1- FUNCTION: THROMBIN, WHICH CLEAVES BONDS AFTER ARG & LYS, CONVERTS  
 CC FIBRINOGEN TO FIBRIN AND ACTIVATES FACTORS V, VII, VIII, XIII,  
 CC AND, IN COMPLEX WITH THROMBOMODULIN, PROTEIN C.  
 CC -1- CATALYTIC ACTIVITY: Preferential cleavage: Arg-I-Gly; activates  
 CC fibrinogen to fibrin and releases fibrinopeptide A and B.  
 CC -1- SUBCELLULAR LOCATION: Extracellular.  
 CC -1- TISSUE SPECIFICITY: SYNTHESIZED IN THE LIVER; FOUND IN PLASMA.  
 CC RESULT FROM THE CARBOXYLATION OF GLUTAMYL RESIDUES BY A MICROSOmal  
 CC ENZYME, THE VITAMIN K-DEPENDENT CARBOXYLASE. THE MODIFIED RESIDUES  
 CC ARE NECESSARY FOR THE CA-DEPENDENT INTERACTION WITH A NEGATIVELY  
 CC CHARGED PHOSPHOLIPID SURFACE, WHICH IS ESSENTIAL FOR THE CONVERSION  
 CC OF PROTHROMBIN TO THROMBIN.  
 CC -1- MISCELLANEOUS: PROTHROMBIN IS ACTIVATED ON THE SURFACE OF A  
 CC PHOSPHOLIPID MEMBRANE THAT BINDS THE AMINO END OF PROTHROMBIN &  
 CC FACTORS VA & XA IN CA-DEPENDENT INTERACTIONS; FACTOR XA REMOVES

CC THE ACTIVATION PEPTIDE & CLEAVES THE REMAINING PART INTO LIGHT &  
 CC HEAVY CHAINS. THE ACTIVATION PROCESS STARTS SLOWLY BECAUSE FACTOR  
 CC V ITSELF HAS TO BE ACTIVATED BY THE INITIAL, SMALL AMOUNTS OF  
 CC THROMBIN.  
 CC -1- MISCELLANEOUS: THROMBIN CAN ITSELF CLEAVE THE AMINO TERMINAL  
 CC FRAGMENT (FRAGMENT 1) OF THE PROTHROMBIN, PRIOR TO ITS ACTIVATION  
 CC BY FACTOR XA.  
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.  
 CC -1- SIMILARITY: CONTAINS 2 KRINGLE DOMAINS.  
 CC -1- DATABASE: NAME-Prozyme technical fact sheet;  
 CC WWW="http://www.prozyme.com/technical/thrombindata.html".  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see http://www.isb-sib.ch/announcement/  
 CC or send an email to license@isb-sib.ch).  
 CC -----  
 CC EMBL; V00135; CAA23451.1; -  
 CC EMBL; J00041; AAA30781.1; -  
 CC PIR; A00915; TBBO  
 CC PIR; S02537; S02537.  
 CC PDB; 1BBR; 31-JAN-94.  
 CC PDB; 1ETR; 31-JAN-94.  
 CC PDB; 1ETS; 31-JAN-94.  
 CC PDB; 1ETT; 31-JAN-94.  
 CC PDB; 1HET; 31-JAN-94.  
 CC PDB; 2PFI; 31-JAN-94.  
 CC PDB; 2PF2; 31-JAN-94.  
 CC PDB; 2SPT; 31-MAY-94.  
 CC PDB; 1MKX; 07-JUL-97.  
 CC PDB; 1MX; 07-JUL-97.  
 CC PDB; 1TBQ; 14-OCT-96.  
 CC PDB; 1TBR; 14-OCT-96.  
 CC PDB; 1TOC; 23-JUL-97.  
 CC PDB; 1VIT; 21-APR-97.  
 CC PDB; 1YCP; 06-MAY-98.  
 CC PDB; 1A0H; 17-JUN-98.  
 CC PDB; 1AVG; 16-FEB-99.  
 CC MEROPS; S01.217; -  
 CC InterPro; IPR001314; Chymotrypsin.  
 CC InterPro; IPR002383; GLA\_blood.  
 CC InterPro; IPR000001; Kringle.  
 CC InterPro; IPR003966; Prothrombin.  
 CC InterPro; IPR001254; Ser\_protease\_Try.  
 CC InterPro; IPR000294; VitK\_dep\_GLA.  
 CC Pfam; PF00051; Kringle; 2.  
 CC Pfam; PF00089; trypsin; 1.  
 CC Pfam; PF00594; gla; 1.  
 CC PRINTS; PR00722; CHYMOTRYPSIN.  
 CC PRINTS; PR00001; GLABLOOD.  
 CC PRINTS; PR00018; KRINGLE.  
 CC PRINTS; PR01505; PROTHROMBIN.  
 CC ProDom; PD000395; Kringle; 2.  
 CC SMART; SM00069; GLA; 1.  
 CC SMART; SM00130; KR; 2.  
 CC SMART; SM00020; Tryp\_SPC; 1.  
 CC PROSITE; PS00011; GLU\_CARBOXYLATION; 1.  
 CC PROSITE; PS00021; KRINGLE\_1; 2.  
 CC PROSITE; PS00070; KRINGLE\_2; 2.  
 CC PROSITE; PS00240; TRYPSIN\_DOM; 1.  
 CC PROSITE; PS00134; TRYPSIN\_HIS; 1.  
 CC PROSITE; PS00135; TRYPSIN\_SER; 1.  
 CC Blood coagulation; Plasma; Calcium-binding; Glycoprotein; Repeat;  
 CC Vitamin K; Zymogen; Gamma-carboxyglutamic acid; Acute phase; Liver;  
 CC Hydrolyase; Serine protease; Kringle; Signal; 3D-structure.  
 CC FT SIGNAL 1 24 POTENTIAL.  
 CC FT PROPEP 1 25 43  
 CC FT CHAIN 44 625.  
 CC FT PEPTIDE 44 199  
 CC FT PEPTIDE 200 317  
 CC FT PEPTIDE 200 317  
 CC ACTIVATION PEPTIDE (FRAGMENT 1).  
 CC ACTIVATION PEPTIDE (FRAGMENT 2).

```
FT CHAIN 318 366 THROMBIN LIGHT CHAIN (A).
FT CHAIN 367 625 THROMBIN HEAVY CHAIN (B).
FT DOMAIN 109 187 KRINGLE 1.
FT DOMAIN 214 292 KRINGLE 2.
FT DOMAIN 367 625 SERINE PROTEASE.
FT SITE 199 200 CLEAVAGE (BY THROMBIN).
FT SITE 317 318 CLEAVAGE (BY FACTOR XA).
FT SITE 366 367 CLEAVAGE (BY FACTOR XA).
FT SITE 409 409 CHARGE RELAY SYSTEM.
FT ACT_SITE 465 465 CHARGE RELAY SYSTEM.
FT ACT_SITE 571 571 CHARGE RELAY SYSTEM.
FT MOD_RES 50 50 GAMMA-CARBOXYGLUTAMIC ACID.
FT MOD_RES 51 51 GAMMA-CARBOXYGLUTAMIC ACID.
FT MOD_RES 58 58 GAMMA-CARBOXYGLUTAMIC ACID.
FT MOD_RES 60 60 GAMMA-CARBOXYGLUTAMIC ACID.
FT MOD_RES 63 63 GAMMA-CARBOXYGLUTAMIC ACID.
FT MOD_RES 64 64 GAMMA-CARBOXYGLUTAMIC ACID.
FT MOD_RES 69 69 GAMMA-CARBOXYGLUTAMIC ACID.
FT MOD_RES 70 70 GAMMA-CARBOXYGLUTAMIC ACID.
FT MOD_RES 73 73 GAMMA-CARBOXYGLUTAMIC ACID.

Query Match 13.7%; Score 357.5; DB 1; Length 625;
Best Local Similarity 29.9%; Pred. No. 3.9e-20;
Matches 92; Conservative 39; Mismatches 114; Indels 63; Gaps 11;

Qy* 172 IPOCSEVETCNGESYRGLMDHTESGKICORWDHQTPHRHKFLPERYPDKGDDNYCRN 231
Db 9 LPGCLALALSLVHSQHVFLAHQAASSLLOR-----ARRANGFLEEVRK- 54

Qy 232 PDGQPRPWCYTLDPHTRWE-----YCAIKTCADNTMNDTDVPLETTE--CIQ 276
Db 55 --GNLERECLE-EPCSRREEAFEALSLSATDAFWAKYFACESARNPREKNECLEGNCAE 111

Qy 277 GQEGYRGTVNTIANGIPQORWDSQYPHEHDMTPENFKCDLRENYCRNPDGS-ESPWCF 335
Db 112 GVGMYRGVSVTRSGIECQLWRSRPHKPEINSTTHPCADLRENFRCRNPDSGITGPWCY 171

Qy 336 TTDPNIRVGYCSQIPNCDSHSGQD-----CYRGNKGNYMGNL 372
Db 172 TTSPTLRRECS-VPVC---GQDRVTVEIPRSGGSTTSQSPLLETCPDRGREYGRGL 226

Qy 373 SOTRSLGTCSMW-DKNMEDLHRHIFWEPDASKLNEYCRNPDDDDAHGPWCYTGPNLPWD 431
Db 227 AVTTSGRCLAWSSEQAKALSQDDFNP-AVPLAENFCRNPDGDEEGAWCYVADQPGDFE 285

Qy 432 YCPISRCE 439
Db 286 YCDLNYCE 293
```

Search completed: April 25, 2003, 15:51:03  
Job time : 32 secs

GenCore version 5.1.4\_p5.4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 25, 2003, 15:48:06 ; Search time 35 seconds  
(without alignments)  
2631.516 Million cell updates/sec

Title: us-09-674-377b-1

Perfect score: 2604

Sequence: 1 ERRRRNTIHEFKSAKTLT.....IPWDYCPISRCGEQTTPTIV 447

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL\_21.\*

1: sp\_archaea.\*  
2: sp\_bacteria.\*  
3: sp\_fungi.\*  
4: sp\_human.\*  
5: sp\_invertebrate.\*  
6: sp\_mammal.\*  
7: sp\_mmc.\*  
8: sp\_organelle.\*  
9: sp\_phase.\*  
10: sp\_plant.\*  
11: sp\_rodent.\*  
12: sp\_virus.\*  
13: sp\_vertebrate.\*  
14: sp\_unclassified.\*  
15: sp\_rvirus.\*  
16: sp\_bacteriap.\*  
17: sp\_archaea.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2452	94.2	728	6 Q9BH09	Q9bh09 felis silve
2	2053	78.8	726	13 Q90978	Q90978 gallus gall
3	1864.5	71.6	710	13 Q91402	Q91402 xenopus. he
4	1471	56.5	290	4 Q02935	Q02935 homo sapien
5	1463	56.2	296	4 Q14519	Q14519 homo sapien
6	1433.5	55.0	285	4 Q8TCE2	Q8tce2 homo sapien
7	1209.5	46.4	704	13 Q90865	Q90865 gallus gall
8	1208	46.4	716	13 Q91691	Q91691 xenopus lae
9	1176	45.2	717	13 P70006	P70006 xenopus lae
10	1172	45.0	709	13 Q902N6	Q902n6 brachydanio
11	1163.5	44.7	567	4 Q13208	Q13208 homo sapien
12	1144	43.9	716	11 P70521	P70521 rattus norv
13	1125	43.2	716	11 Q91XG8	Q91xg8 mus musculu
14	1085.5	41.7	812	11 Q9ROW3	Q9row3 rattus norv
15	1080.5	41.5	806	6 Q18793	Q18793 macropus eu
16	1064	40.9	812	11 Q91WJ5	Q91wj5 mus musculu

17 1042 40.0 810 4 Q15146  
18 1040 39.9 215 13 Q42341  
19 985 37.8 359 6 Q8WMR1  
20 978 37.6 208 4 Q9BYM0  
21 978 37.6 210 4 Q13494  
22 966.5 37.1 648 4 Q9H1V4  
23 943.5 36.2 2869 6 Q28398  
24 896 34.4 211 11 Q55027  
25 615 23.6 109 6 Q9N1B8  
26 444.5 17.1 454 6 Q46506  
27 439 16.9 75 6 Q9BGN9  
28 387.5 14.9 948 5 Q9Y1V6  
29 335.5 12.9 313 13 Q9P078  
30 326 12.5 607 13 Q91001  
31 320.5 12.3 608 13 Q9PTW7  
32 297.5 11.4 559 11 Q91VP2  
33 294 11.3 516 4 Q9BU99  
34 277 10.6 334 6 Q46507  
35 274 10.5 562 6 Q8SQ23  
36 273 10.5 132 4 Q16609  
37 272 10.4 113 4 Q9UIR5  
38 260 10.0 113 4 Q9UIR6  
39 256 9.8 113 4 Q9UIR7  
40 249 9.6 105 4 Q9UIR8  
41 244.5 9.4 385 5 Q25101  
42 242 9.3 145 6 Q28911  
43 235.5 9.0 946 13 Q07153  
44 218 8.4 393 4 Q9BRB6  
45 218 8.4 724 5 Q9V6K3

#### ALIGNMENTS

#### RESULT 1

Q9BH09 ID Q9BH09 PRELIMINARY; PRT; 728 AA.  
AC Q9BH09;  
DT 01-JUN-2001 (TREMBLrel. 17, Created)  
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)  
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
DE Hepatocyte growth factor HGF.  
OS Felis silvestris catus (Cat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Carnivora; Fissipedia; Felidae; Felis.  
OX NCBI\_TaxID=9685;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=LIVER;  
RA Kobayashi Y., Nakamura N., Ishizaka T., Masuda K., Ohno K.,  
RA Tsujimoto H.;  
RT "Molecular cloning of feline hepatocyte growth factor (HGF) cDNA."  
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.  
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE  
CC TRYPSIN FAMILY.  
DR EMBL; AB046610; BAB21499.1; -.  
DR HSSP; P14210; 1BHT.  
DR MEROPS; S01.982; -.  
DR InterPro; IPR001314; Chymotrypsin.  
DR InterPro; IPR000001; Kringle.  
DR InterPro; IPR003014; PAN.  
DR InterPro; IPR003609; Pan\_app.  
DR InterPro; IPR001254; Ser\_protease\_Try.  
DR Pfam; PF00051; kringle; 4.  
DR Pfam; PF00024; PAN; 1.  
DR Pfam; PF00089; trypsin; 1.  
DR PRINTS; PR00722; CHYMOTRYPSIN.  
DR PRINTS; PR00018; KRINGLE.  
DR ProDom; PD000395; Kringle; 4.  
DR SMART; SM00130; KR; 4.  
DR SMART; SM00473; PAN\_AP; 1.  
DR SMART; SM00020; TRY\_P-SPC; 1.  
DR PROSITE; PS00021; KRINGLE\_1; UNKNOWN\_4.

CC	PROSITE: PS50070; KRINGLE_2; 4.	CC	-1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE TRYPSIN FAMILY.
DR	PROSITE: PS50240; TRYPSIN_DOM; 1.	DR	EMBL: X84045; CAA58964.1; -.
KW	Hydrolase; Serine protease.	DR	EMBL: X80131; CAA56430.1; -.
SQ	SEQUENCE 728 AA; 83067 MW; 8D7FAA333D1E190A CRC64;	DR	HSSP: P14210; 1BHT.
	Query Match 94.2%; Score 2452; DB 6; Length 728;	DR	InterPro: IPR001314; Chymotrypsin.
	Best Local Similarity 92.6%; Pred. No. 1.2e-219;	DR	InterPro: IPR000001; Kringle.
	Matches 414; Conservative 21; Mismatches 12; Indels 0; Gaps 0;	DR	InterPro: IPR003014; PAN.
QY	1 EKRRNTIHEFKKSAKTLIKIDPALKIKTKKVNADOCANCRTRNKGLPFTCKAFVDEK 60	DR	InterPro: IPR001254; Ser. protease_Try.
Db	1 :     :     :     :     :     :     :     :     :     :	DR	InterPro: IPR001309; Pan_app.
	:     :     :     :     :     :     :     :     :	DR	InterPro: IPR001254; Ser. protease_Try.
QY	61 ARKQCLWFFPNSMSSGVKKEGHEFDLYENKDYIRNCIIIGKRSYKGTIVTSIKSGIKCOP 120	DR	InterPro: IPR001254; Ser. protease_Try.
Db	1 :     :     :     :     :     :     :     :     :	DR	InterPro: IPR001254; Ser. protease_Try.
	:     :     :     :     :     :     :     :     :	DR	InterPro: IPR001254; Ser. protease_Try.
QY	90 ARKRCWFFPNSMSSGVKKEGHEFDLYENKDYIRNCIIIGKRSYKGTIVTSIKSGIKCOP 149	DR	InterPro: IPR001254; Ser. protease_Try.
Db	1 :     :     :     :     :     :     :     :     :	DR	InterPro: IPR001254; Ser. protease_Try.
	:     :     :     :     :     :     :     :     :	DR	InterPro: IPR001254; Ser. protease_Try.
QY	121 WSMIPIHEHSLPSSYRGKDLQENYCRNPRGEGGPGWCFSTNPEVRYEVCIDIPQCEVEEC 180	DR	InterPro: IPR001254; Ser. protease_Try.
Db	1 :     :     :     :     :     :     :     :     :	DR	InterPro: IPR001254; Ser. protease_Try.
	:     :     :     :     :     :     :     :     :	DR	InterPro: IPR001254; Ser. protease_Try.
QY	150 WSMIPIHEHSLPSSYRGKDLQENYCRNPRGEGGPGWCFSTNPEVRYEVCIDIPQCEVEEC 209	DR	InterPro: IPR001254; Ser. protease_Try.
Db	1 :     :     :     :     :     :     :     :     :	DR	InterPro: IPR001254; Ser. protease_Try.
	:     :     :     :     :     :     :     :     :	DR	InterPro: IPR001254; Ser. protease_Try.
QY	181 MTCNGESYRGKDLQENYCRNPRGEGGPGWCFSTNPEVRYEVCIDIPQCEVEEC 240	DR	InterPro: IPR001254; Ser. protease_Try.
Db	1 :     :     :     :     :     :     :     :     :	DR	InterPro: IPR001254; Ser. protease_Try.
	:     :     :     :     :     :     :     :     :	DR	InterPro: IPR001254; Ser. protease_Try.
QY	210 MTCNGESYRGKDLQENYCRNPRGEGGPGWCFSTNPEVRYEVCIDIPQCEVEEC 269	DR	InterPro: IPR001254; Ser. protease_Try.
Db	1 :     :     :     :     :     :     :     :     :	DR	InterPro: IPR001254; Ser. protease_Try.
	:     :     :     :     :     :     :     :     :	DR	InterPro: IPR001254; Ser. protease_Try.
QY	241 YTLDPHTRWEYCAIKTCADNTMDTDVPLETTECIQOGEGYRGTVNTIWNIGPCQRWDS 300	DR	InterPro: IPR001254; Ser. protease_Try.
Db	1 :     :     :     :     :     :     :     :     :	DR	InterPro: IPR001254; Ser. protease_Try.
	:     :     :     :     :     :     :     :     :	DR	InterPro: IPR001254; Ser. protease_Try.
QY	270 YTLDPHTRWEYCAIKTCADNTMDTDVPLETTECIQOGEGYRGTVNTIWNIGPCQRWDS 329	DR	InterPro: IPR001254; Ser. protease_Try.
Db	1 :     :     :     :     :     :     :     :     :	DR	InterPro: IPR001254; Ser. protease_Try.
	:     :     :     :     :     :     :     :     :	DR	InterPro: IPR001254; Ser. protease_Try.
QY	301 QYHEHDMTPENFKCKDLRENYCRNPRGEGGPGWCFSTNPEVRYEVCIDIPQCEVEEC 360	DR	InterPro: IPR001254; Ser. protease_Try.
Db	1 :     :     :     :     :     :     :     :     :	DR	InterPro: IPR001254; Ser. protease_Try.
	:     :     :     :     :     :     :     :     :	DR	InterPro: IPR001254; Ser. protease_Try.
QY	330 QYPHOHDIITENPKCKDLRENYCRNPRGEGGPGWCFSTNPEVRYEVCIDIPQCEVEEC 389	DR	InterPro: IPR001254; Ser. protease_Try.
Db	1 :     :     :     :     :     :     :     :     :	DR	InterPro: IPR001254; Ser. protease_Try.
	:     :     :     :     :     :     :     :     :	DR	InterPro: IPR001254; Ser. protease_Try.
QY	361 YRGNGKMYGNLSOTRSGLTCSMDKNNMEDLHRHIFWEPDASKLNENYCRNPDHAGPW 420	DR	InterPro: IPR001254; Ser. protease_Try.
Db	1 :     :     :     :     :     :     :     :     :	DR	InterPro: IPR001254; Ser. protease_Try.
	:     :     :     :     :     :     :     :     :	DR	InterPro: IPR001254; Ser. protease_Try.
QY	390 YRGNGKMYGNLSOTRSGLTCSMDKNNMEDLHRHIFWEPDASKLNENYCRNPDHAGPW 449	DR	InterPro: IPR001254; Ser. protease_Try.
Db	1 :     :     :     :     :     :     :     :     :	DR	InterPro: IPR001254; Ser. protease_Try.
	:     :     :     :     :     :     :     :     :	DR	InterPro: IPR001254; Ser. protease_Try.
QY	421 CYTGNPLIPWDYCPISRCGDTTPTIV 447	DR	InterPro: IPR001254; Ser. protease_Try.
Db	1 :     :     :     :     :     :     :     :     :	DR	InterPro: IPR001254; Ser. protease_Try.
	:     :     :     :     :     :     :     :     :	DR	InterPro: IPR001254; Ser. protease_Try.
QY	450 CYTGNPLIPWDYCPISRCGDTTPTIV 476	DR	InterPro: IPR001254; Ser. protease_Try.
Db	1 :     :     :     :     :     :     :     :     :	DR	InterPro: IPR001254; Ser. protease_Try.
	:     :     :     :     :     :     :     :     :	DR	InterPro: IPR001254; Ser. protease_Try.
RESULT 2		DR	InterPro: IPR001254; Ser. protease_Try.
Q90978	PRELIMINARY; PRT; 726 AA.	DR	InterPro: IPR001254; Ser. protease_Try.
ID	Q90978; Q90966;	DR	InterPro: IPR001254; Ser. protease_Try.
AC	01-JAN-1998 (TrEMBLrel. 05, Created)	DR	InterPro: IPR001254; Ser. protease_Try.
DT	01-JAN-1998 (TrEMBLrel. 05, Last sequence update)	DR	InterPro: IPR001254; Ser. protease_Try.
DT	01-MAR-2002 (TrEMBLrel. 20, Last annotation update)	DR	InterPro: IPR001254; Ser. protease_Try.
DE	Hepatocyte growth factor / scatter factor precursor.	DR	InterPro: IPR001254; Ser. protease_Try.
O5	Gallus gallus (Chicken).	DR	InterPro: IPR001254; Ser. protease_Try.
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	DR	InterPro: IPR001254; Ser. protease_Try.
OC	Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;	DR	InterPro: IPR001254; Ser. protease_Try.
OC	Gallus...	DR	InterPro: IPR001254; Ser. protease_Try.
OX	NCBI_TaxID=9031;	DR	InterPro: IPR001254; Ser. protease_Try.
RN	[1]	DR	InterPro: IPR001254; Ser. protease_Try.
RP	SEQUENCE FROM N.A.	DR	InterPro: IPR001254; Ser. protease_Try.
RC	SEQUENCE-EMBRYO;	DR	InterPro: IPR001254; Ser. protease_Try.
RX	MEDLINE=96029010; PubMed=7554499;	DR	InterPro: IPR001254; Ser. protease_Try.
RA	Thery C., Sharpe M.J., Batley S.J., Stern C.D., Gherardi E.;	DR	InterPro: IPR001254; Ser. protease_Try.
RT	"Expression of HGF/SF, HGF/SF, and c-met suggests new functions	DR	InterPro: IPR001254; Ser. protease_Try.
RT	during early chick development."	DR	InterPro: IPR001254; Ser. protease_Try.
RL	Dev. Genet. 17:90-101(1995).	DR	InterPro: IPR001254; Ser. protease_Try.
RN	[2]	DR	InterPro: IPR001254; Ser. protease_Try.
RP	SEQUENCE OF 1-409 FROM N.A.	DR	InterPro: IPR001254; Ser. protease_Try.
RC	STRAIN-WHITE LEHORN, AND RHODE ISLAND RED X LIGHT SUSSEX;	DR	InterPro: IPR001254; Ser. protease_Try.
RX	TISSUE-EMBRYO;	DR	InterPro: IPR001254; Ser. protease_Try.
RA	MEDLINE=95237013; PubMed=7720585;	DR	InterPro: IPR001254; Ser. protease_Try.
RT	Sharpe M.J., Gherardi E.;	DR	InterPro: IPR001254; Ser. protease_Try.
RT	"A role for HGF/SF in neural induction and its expression in Hensen's	DR	InterPro: IPR001254; Ser. protease_Try.
RT	node during gastrulation."	DR	InterPro: IPR001254; Ser. protease_Try.
RL	Development 121:813-824(1995).	DR	InterPro: IPR001254; Ser. protease_Try.



```

GN HGF.
OS Xenopus.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
OC Xenopodinae.
OX NCBI_TaxID=8353;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=TAILBUD;
RX MEDLINE=95267690; PubMed=7748783;
RA Nakamura H., Tashiro K., Nakamura T., Shiohawa K.;
RT "Molecular cloning of Xenopus HGF cDNA and its expression studies in
RL xenopus early embryogenesis.";
RL Mech. Dev. 49:123-131(1995).
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC TRYPSIN FAMILY.
DR EMBL; S77422; AAB34354.2;
DR HSP; P14210; IBHT.
DR MEROPS; S01.976;
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR003014; PAN.
DR InterPro; IPR003609; Pan_app.
DR InterPro; IPR001254; Ser.protease_Try.
DR Pfam; PF00051; kringle; 4.
DR Pfam; PF00024; PAN; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 4.
DR SMART; SM00130; KR; 4.
DR SMART; SM00473; PAN_AP; 1.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS00021; KRINGLE_1; 3.
DR PROSITE; PS00070; KRINGLE_2; 4.
DR PROSITE; PS0240; TRYPSIN_DOM; 1.
KW Hydrolase; Serine protease.
SQ SEQUENCE 710 AA; 81487 MW; 5FE6480BE31C27FC CRC64;
Query Match 71.6%; Score 1864.5; DB 13; Length 710;
Best Local Similarity 69.3%; Pred. No. 5.9e-165;
Matches 305; Conservative 64; Mismatches 58; Indels 3; Gaps 1;

Qy 3 KRRTITHEPKSAKTLIKIDPAKIKTKYNTADQCANRTRNKGLPFTCKAFVFDKAR 62
Db 21 KRNAFDYKTAETTLRLNKALEVKTKMFTTENCACRCSRNKGLPFTCKAFADKNI 80

Qy 63 KQCLWFPNSMSSGVKKEFGHEFDLYENKDYIRNCIIIGKRSYKGTYSITKSGIKCQPS 122
Db* 81 KRCHWFSNTMSAGIKDKYIDISFDLYEKKYIRDCIHGKSGYRGTNRVTKRGLACQPN 140

Qy 123 SMIPHEHSFLPSSYRGKDLQENYCRNPRGEGGPGWCFTSNPEVRYEVCIDIPOCSEVECM 182
Db 141 SMIPHEHSFLPSTYRGKDLKENYCRNPKGESGPGWCFTKSPVHRDVCIDIPOCSEVDCV 200

Qy 183 CNGESYRGLMDHTSGKTCQRWDHQTTHRHKFLPERYPDKGFDNDCYRNPQGRPWYCT 242
Db 201 CNGEYRGPMDYTESGKRCQWDQLQRPKHKFLPERYPKNGKGLNDCYRNPQGRPWYCT 260

Qy 243 LDHTRWEYCAIKCADMTMDTDVPLETTECIQGGYRGYVNTIWNIGIPCORWDSQY 302
Db 261 LDPTDSWEFCAIKPCVHSIVANTDI---TKDCMKGGGEGYRGVSTTYNGIQCRWDSQF 317

Qy 303 PHEHDMTPENFKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCMDSHGQDCYR 362
Db 318 PHLHNFTEPNYKCDLSENYCRNPDGSESPWCFTTDPNIRIGHCSQIKKCOASNQECY 377

Qy 363 GNGKNGMNLQTSRGLTCSMWDKNMEDLHRHIFWEPDASKLNENYCRNPDDAHGPWCY 422
Db 378 GNGSYKGTLSRTRFLRFLPCSMWENLQDLKRHTNEPDSVILQKNYCRNPDDAHGPWCY 437

Qy 423 TGNPLIPWDYCPISRCEDGT 442
Db 442 TGNPLIPWDYCPISRCEDGT 442

```

```

Db 438 TDDPFPVPMWYCPISRCEDGT 457

RESULT 4
Q02935
ID Q02935 PRELIMINARY; PRT; 290 AA.
AC Q02935;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Hepatocyte growth factor, heavy chain precursor.
GN HGF.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=PLACENTA;
RX MEDLINE=91200041; PubMed=1826653;
RA Miyazawa K., Kitamura A., Naka D., Kitamura N.;
RT "An alternatively processed mRNA generated from human hepatocyte
RT growth factor gene.";
RL Eur. J. Biochem. 197:15-22(1991).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=93087571; PubMed=1280830;
RA Hartmann G., Naldini L., Weidner K.M., Sachs M., Vigna E.,
RA Comoglio P.M., Birdmeier W.;
RT "A functional domain in the heavy chain of scatter factor/hepatocyte
RT growth factor binds the c-Met receptor and induces cell dissociation
RT but not mitogenesis.";
RL Proc. Natl. Acad. Sci. U.S.A. 89:11574-11578(1992).
CC -1- FUNCTION: HGF IS A POTENT MITOGEN FOR MATURE PARENCHYMAL
CC HEPATOCYTE CELLS, SEEMS TO BE AN HEPATOTROPHIC FACTOR, AND ACTS AS
CC GROWTH FACTOR FOR A BROAD SPECTRUM OF TISSUES AND CELL TYPES. IT
CC HAS NO DETECTABLE PROTEASE ACTIVITY.
CC -1- SIMILARITY: CONTAINS TWO KRINGLE REGIONS.
CC -1- SIMILARITY: HIGH, TO OTHER HGF; LOWER, TO PLASMINOGEN.
DR EMBL; X57574; CAA04802.1;
DR EMBL; L02931; AAA52649.1;
DR HSP; P14210; IBHT.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR003014; PAN.
DR InterPro; IPR003609; Pan_app.
DR Pfam; PF00051; kringle; 2.
DR Pfam; PF00024; PAN; 1.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 2.
DR SMART; SM00130; KR; 2.
DR SMART; SM00473; PAN_AP; 1.
DR PROSITE; PS00021; KRINGLE_1; 2.
DR PROSITE; PS00070; KRINGLE_2; 2.
KW Growth factor; Kringle; Signal.
FT SIGNAL 1 31
FT CHAIN 32 290
FT DOMAIN 128 206
FT DOMAIN 211 288
FT MOD_RES 32 32
FT (BY SIMILARITY).
SQ SEQUENCE 290 AA; 33765 MW; C8A18A6F0D63200A CRC64;
Query Match 56.5%; Score 1471; DB 4; Length 290;
Best Local Similarity 99.6%; Pred. No. 8.3e-129;
Matches 256; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ERKRNTTHEPKSAKTLIKIDPAKIKTKYNTADQCANRTRNKGLPFTCKAFVFDK 60
Db 32 QRKRNTTHEPKSAKTLIKIDPAKIKTKYNTADQCANRTRNKGLPFTCKAFVFDK 91

Qy 61 ARKQCLWFPNSMSSGVKKEFGHEFDLYENKDYIRNCIIIGKRSYKGTYSITKSGIKCQ 120
Db 92 ARKQCLWFPNSMSSGVKKEFGHEFDLYENKDYIRNCIIIGKRSYKGTYSITKSGIKCQ 151

```

```

QY 121 WSSMIPHEHSLPSSYRGKDLQENYCRNPRGEGGPGWCFTSNPEVRYEVCDDIPQCSEVEC 180
DB 152 WSSMIPHEHSLPSSYRGKDLQENYCRNPRGEGGPGWCFTSNPEVRYEVCDDIPQCSEVEC 211
QY 181 MTCNGESYRGLMDHTESGKICQWRDHTQPHRHKFLPERYPDKGFDNNYCRNPDGQPRPWC 240
DB 212 MTCNGESYRGLMDHTESGKICQWRDHTQPHRHKFLPERYPDKGFDNNYCRNPDGQPRPWC 271
QY 241 YTLDPHTRWEYCAIKTC 257
DB 272 YTLDPHTRWEYCAIKTC 288

RESULT 5
Q14519
ID Q14519 PRELIMINARY; PRT; 296 AA.
AC Q14519;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Competitive HGF antagonist.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheraia; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LUNG;
RA Chan A.M.-L., Rubin J.S., Bottaro D.P., Hirschfield D.W., Chedid M.,
RA Aaronson S.A.;
RT "Identification of a Competitive HGF Antagonist Encoded by an
RT Alternative Transcript.";
RL Science 0:0-0(1991):80.1; -.
DR EMBL; M77227; AAA35980.1; -.
DR HSSP; P14210; 1BHT.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR003014; PAN.
DR InterPro; IPR003609; Pan_app.
DR Pfam; PF00051; kringle; 2.
DR Pfam; PF00024; PAN; 1.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 2.
DR SMART; SM00130; KR; 2.
DR SMART; SM00473; PAN_AP; 1.
DR PROSITE; PS00021; KRINGLE_1; 2.
DR PROSITE; PS00070; KRINGLE_2; 2.
SQ SEQUENCE 296 AA; 34546 MW; A45E456B87AE03BE CRC64;

Query Match 56.2%; Score 1463; DB 4; Length 296;
Best Local Similarity 98.1%; Pred. No. 4.7e-128;
Matches 256; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 ERKRRTIHEFKSAKTTLIKIDPALKIKTKKVTADQCANRCTRNKGLPFTCKAFVFDK 60
DB 32 QKRRNTIHEFKSAKTTLIKIDPALKIKTKKVTADQCANRCTRNKGLPFTCKAFVFDK 91
QY 61 ARKQCLWFPNMSGSGVKKKEGHEFDLYENKDYIRNCIIIGKGRSYKGTVSITKSGIKCQP 120
DB 92 ARKQCLWFPNMSGSGVKKKEGHEFDLYENKDYIRNCIIIGKGRSYKGTVSITKSGIKCQP 151
QY 121 WSSMIPHEHSLPSSYRGKDLQENYCRNPRGEGGPGWCFTSNPEVRYEVCDDIPQCSEVEC 180
DB 152 WSSMIPHEHSLPSSYRGKDLQENYCRNPRGEGGPGWCFTSNPEVRYEVCDDIPQCSEVEC 206
QY 181 MTCNGESYRGLMDHTESGKICQWRDHTQPHRHKFLPERYPDKGFDNNYCRNPDGQPRPWC 240
DB 207 MTCNGESYRGLMDHTESGKICQWRDHTQPHRHKFLPERYPDKGFDNNYCRNPDGQPRPWC 266
QY 241 YTLDPHTRWEYCAIKTC 257
DB 267 YTLDPHTRWEYCAIKTC 283

RESULT 7
Q90865
ID Q90865 PRELIMINARY; PRT; 704 AA.
AC Q90865;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Hepatocyte growth factor-like/macrophage stimulating protein.
GN HGFI/MSP.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LIVER;
RX MEDLINE=96029010; PubMed=7554499;
RA Thery C., Sharpe M.J., Batley S.J., Stern C.D., Gherardi E.;
RT "Expression of HGFI/MSP, HGFI/MSP and c-met suggests new functions
RT during early chick development.";
RL Dev. Genet. 17:90-101(1995).
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC TRYPsin FAMILY.
CC EMBL; X84043; CAA58862.1; -.
DR HSSP; P00747; ICEA.
```

```

RESULT 6
Q8TCE2
ID Q8TCE2 PRELIMINARY; PRT; 285 AA.
AC Q8TCE2;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Similar to hepatocyte growth factor (hepapoietin A, scatter
DE factor).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheraia; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=PLACENTA;
RA Strausberg R.;
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC022308; AAH22308.1; -.
SQ SEQUENCE 285 AA; 33234 MW; 0A93B073EA86EA61 CRC64;

Query Match 55.0%; Score 1433.5; DB 4; Length 285;
Best Local Similarity 97.7%; Pred. No. 2.5e-125;
Matches 251; Conservative 1; Mismatches 0; Indels 5; Gaps 1;

QY 1 ERKRRTIHEFKSAKTTLIKIDPALKIKTKKVTADQCANRCTRNKGLPFTCKAFVFDK 60
DB 32 QKRRNTIHEFKSAKTTLIKIDPALKIKTKKVTADQCANRCTRNKGLPFTCKAFVFDK 91
QY 61 ARKQCLWFPNMSGSGVKKKEGHEFDLYENKDYIRNCIIIGKGRSYKGTVSITKSGIKCQP 120
DB 92 ARKQCLWFPNMSGSGVKKKEGHEFDLYENKDYIRNCIIIGKGRSYKGTVSITKSGIKCQP 151
QY 121 WSSMIPHEHSLPSSYRGKDLQENYCRNPRGEGGPGWCFTSNPEVRYEVCDDIPQCSEVEC 180
DB 152 WSSMIPHEHSLPSSYRGKDLQENYCRNPRGEGGPGWCFTSNPEVRYEVCDDIPQCSEVEC 206
QY 181 MTCNGESYRGLMDHTESGKICQWRDHTQPHRHKFLPERYPDKGFDNNYCRNPDGQPRPWC 240
DB 207 MTCNGESYRGLMDHTESGKICQWRDHTQPHRHKFLPERYPDKGFDNNYCRNPDGQPRPWC 266
QY 241 YTLDPHTRWEYCAIKTC 257
DB 267 YTLDPHTRWEYCAIKTC 283

RESULT 7
Q90865
ID Q90865 PRELIMINARY; PRT; 704 AA.
AC Q90865;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Hepatocyte growth factor-like/macrophage stimulating protein.
GN HGFI/MSP.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LIVER;
RX MEDLINE=96029010; PubMed=7554499;
RA Thery C., Sharpe M.J., Batley S.J., Stern C.D., Gherardi E.;
RT "Expression of HGFI/MSP, HGFI/MSP and c-met suggests new functions
RT during early chick development.";
RL Dev. Genet. 17:90-101(1995).
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC TRYPsin FAMILY.
CC EMBL; X84043; CAA58862.1; -.
DR HSSP; P00747; ICEA.
```

DR MEROPS: S01.977; --  
 DR InterPro: IPR001314; Chymotrypsin.  
 DR InterPro: IPR000001; Kringle.  
 DR InterPro: IPR003014; PAN.  
 DR InterPro: IPR003609; Pan\_app.  
 DR InterPro: IPR001254; Ser\_protease\_Try.  
 DR Pfam: PF00051; kringle; 4.  
 DR Pfam: PF00024; PAN; 1.  
 DR Pfam: PF00089; trypsin; 1.  
 DR PRINTS: PR00018; KRINGLE.  
 DR ProDom: PD000395; Kringle; 4.  
 DR SMART: SM00130; KR; 4.  
 DR SMART: SM00473; PAN\_AP; 1.  
 DR SMART: SM00020; Tryp\_SPC; 1.  
 DR PROSITE: PS00021; KRINGLE\_1; 4.  
 DR PROSITE: PS00070; KRINGLE\_2; 4.  
 DR PROSITE: PS0240; TRYPSIN\_DOM; 1.  
 KW Hydroxylase; Serine protease.  
 SQ SEQUENCE 704 AA; 79341 MW; CAB0D8CC41367C37 CRC64;  
  
 Query Match 46.4%; Score 1209.5; DB 13; Length 704;  
 Best Local Similarity 46.0%; Pred. No. 5.6e-104;  
 Matches 207; Conservative 71; Mismatches 149; Indels 23; Gaps 8;  
  
 QY 5 RNTIHEFKSAKTLIKIDPALKIKTKYNTADQCANRCTRNNKGLPFTCKAFVFDKARKQ 64  
 DB 20 RSLPNDQFRLRGTETELRAAPNEPPSAPAHGAQAQCAQRCANRP-----DCRAFHERQSOL 75  
 QY 65 CLWFPNSMSSGVKKEFGHEFDLYENKDIYRNCIIIGKGRSYKGTVSITKSGIKCQPSWSM 124  
 DB 76 COLLPWSQRSQARLQKNTHYDLYQKDLRECIYANGSYRGTTRTTERGLRQCQHWQAT 135  
 QY 125 IPHEHSFLPSSYRGKDLQENYCRNPRGEGGPMWFTSNPEVRYEVCIDIPQCSVEVCMTCN 184  
 DB 136 TPDHREFLSLRG--LEENYCRNPRDRGRGPCWYVDNVRHQSGIKKCEDAVCMTCN 193  
 QY 185 GESYRGMDHTESGKTCQRMWDHTPHRHKFLPERYPDKGFDNNYCRNPDGQPRPWCYTLD 244  
 DB 194 GEDYRGFVDHTESGTCQRMWDLQHPHKHPYHPDKYPEKGLDNNYCRNPDSSSEQPMWCYTLD 253  
 QY 245 PHTRWYCAIKCADTMDNDTVPLE--TTECIQGGEGYRGVTNTIWNIGIPCQWRDWSQYP 303  
 DB 254 PALEREFCIRVCKKPR-----PINVTTCYRGKGGYRGVNTVTSIGIPCQWRDQAQTL 308  
 QY 304 HEHDMTPENFKCKDLRENYCRNPDGSESPWCFTTDNIRVGYCSQIPNCDMS--HGQDCYR 362  
 DB 309 HRHHFVPSKYPCKDLQENYCRNPDGSEAPWCFTTRGMRVAFCFHRRCDDELDAECYH 368  
 QY 363 GNGKNTWGNLSQTRSGLTCSMDKNMEDLHRHI-----FWEPDASKLNNYCRNPDODAH 417  
 DB 369 GHGRYGHVSKTRKGIQTCQRMWDATTP----HVPQISPTTHPEA--HLEKNYCRNPDNDSH 423  
 QY 418 GPMCKTGNLIPWDCYPSRCEGDTTPTIV 447  
 DB 424 GPMCKTGNLIPWDCYPSRCEGDTTPTIV 453  
  
 RESULT 8  
 Q91691 PRELIMINARY; PRT; 716 AA.  
 ID Q91691  
 AC Q91691;  
 DT 01-NOV-1996 (Tremblrel. 01, Created)  
 DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)  
 DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)  
 DE Growth factor Livertine.  
 OS Xenopus laevis (African clawed frog).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;  
 OC Xenopodinae; Xenopus.  
 OX NCBI\_TaxID=8355;  
 RN [1]  
 RP SEQUENCE FROM N.A.

RA Ruiz i Altaba A., Thery C.;  
 RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: U57455; AAB52574.1;  
 DR HSP: P00747; ICEA.  
 DR InterPro: IPR000001; Kringle.  
 DR InterPro: IPR003014; PAN.  
 DR InterPro: IPR003609; Pan\_app.  
 DR InterPro: IPR001254; Ser\_protease\_Try.  
 DR Pfam: PF00051; kringle; 4.  
 DR Pfam: PF00024; PAN; 1.  
 DR Pfam: PF00089; trypsin; 1.  
 DR PRINTS: PR00018; KRINGLE.  
 DR ProDom: PD000395; Kringle; 4.  
 DR SMART: SM00130; KR; 4.  
 DR SMART: SM00473; PAN\_AP; 1.  
 DR SMART: SM00020; Tryp\_SPC; 1.  
 DR PROSITE: PS00021; KRINGLE\_1; 4.  
 DR PROSITE: PS00070; KRINGLE\_2; 4.  
 DR PROSITE: PS0240; TRYPSIN\_DOM; 1.  
 KW Hydroxylase; Serine protease.  
 SQ SEQUENCE 716 AA; 81971 MW; 508376A0E4398798 CRC64;  
  
 Query Match 46.4%; Score 1208; DB 13; Length 716;  
 Best Local Similarity 46.7%; Pred. No. 7.9e-104;  
 Matches 207; Conservative 75; Mismatches 149; Indels 12; Gaps 6;  
  
 QY 5 RNTIHEFKSAKTLIKIDPALKIKTKYNTADQCANRCTRNNKGLPFTCKAFVFDKARKQ 64  
 DB 31 RSLNDYQSRKGLVHMNG--GVKQEIQIQCAKQCS-----LLDCRSFYNNKKSQT 85  
 QY 65 CLWFPNSMSSGVKKEFGHEFDLYENKDIYRNCIIIGKGRSYKGTVSITKSGIKCQPSWSM 124  
 DB 86 CRLLPWTQNSANVLLQRYDLYQKDYRDCVAGNGTYRGTVTSKSGRTCORMLK 145  
 QY 125 IPHEHSFLPSSYRGKDLQENYCRNPRGEGGPMWFTSNPEVRYEVCIDIPQCSVEVCMTCN 184  
 DB 146 FPDHREFLSLRG--PELENYCRNPDSDPEGPMWYTTDNIRHGYCGIKKCEDAVCLTCN 203  
 QY 185 GESYRGMDHTESGKTCQRMWDHTPHRHKFLPERYPDKGFDNNYCRNPDGQPRPWCYTLD 244  
 DB 204 GEDYRGFVDHTESGTCQRMWDLQHPHKHPYHPDKYPEKGLDNNYCRNPDSSSEQPMWCYTLD 263  
 QY 245 PHTRWYCAIKCADTMDNDTVPLE--TTECIQGGEGYRGVTNTIWNIGIPCQWRDWSQYP 304  
 DB 264 PNYEFCIRVCKKPR-----PINVTTCYRGKGGYRGVNTVTSIGIPCQWRDQAQTL 320  
 QY 305 HEHDMTPENFKCKDLRENYCRNPDGSESPWCFTTDNIRVGYCSQIPNCDMS--HGQDCYR 363  
 DB 321 SHRFLPEKYPCKDLQENYCRNPDGSEAPWCFTTRGMRVAFCFHRRCDDELDAECYH 380  
 QY 364 GNGKNTWGNLSQTRSGLTCSMDKNMEDLHRHI-----FWEPDASKLNNYCRNPDODAH 423  
 DB 381 NGELYSGYVSKTRKGIQTCQRMWDATTP----HVPQISPTTHPEA--HLEKNYCRNPDNDSH 439  
 QY 424 GPMCKTGNLIPWDCYPSRCEGDTTPTIV 446  
 DB 440 GPMCKTGNLIPWDCYPSRCEGDTTPTIV 462  
  
 RESULT 9  
 P70006 PRELIMINARY; PRT; 717 AA.  
 ID P70006  
 AC P70006;  
 DT 01-FEB-1997 (Tremblrel. 02, Created)  
 DT 01-FEB-1997 (Tremblrel. 02, Last sequence update)  
 DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)  
 DE Hepatocyte growth factor-like protein precursor.  
 OS Xenopus laevis (African clawed frog).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;  
 OC Xenopodinae; Xenopus.  
 OX NCBI\_TaxID=8355;  
 RN [1]

RP SEQUENCE FROM N.A.  
RC TISSUE-KIDNEY;  
RX MEDLINE=96404125; PubMed=8808403;  
RA Aberger F., Schmidt G., Richter K.;  
RT "The xenopus homologue of hepatocyte growth factor-like protein is  
RT specifically expressed in the presumptive neural plate during  
RT gastrulation.";  
RL Mech. Dev. 54:23-37(1996).  
DR EMBL; Y08734; CAA69989.1; .  
DR HSSP; P00747; ICEA.  
DR MEROPS; S01.977;  
DR InterPro; IPR000001; Kringle.  
DR InterPro; IPR003014; PAN.  
DR InterPro; IPR003609; Pan\_app.  
DR InterPro; IPR001254; Ser.protease\_Try.  
DR Pfam; PF00051; kringle; 4.  
DR Pfam; PF00024; PAN; 1.  
DR Pfam; PF00089; trypsin; 1.  
DR PRINTS; PR00018; KRINGLE.  
DR ProDom; PD000395; Kringle; 4.  
DR SMART; SM00130; KR; 4.  
DR SMART; SM00473; PAN\_AP; 1.  
DR SMART; SM00020; Tryp\_SPC; 1.  
DR PROSITE; PS00021; KRINGLE\_1; 3.  
DR PROSITE; PS50070; KRINGLE\_2; 4.  
DR PROSITE; PS50240; TRYPsin\_DOM; 1.  
KW Hydrolase; Serine protease; Signal.  
FT SIGNAL 1 28 POTENTIAL.  
FT CHAIN 29 717 HEPATOCYTE GROWTH FACTOR-LIKE PROTEIN.  
SQ SEQUENCE 717 AA; 82017 MW; 6F877A432C8CDD54 CRC64;

Query Match 45.2%; Score 1176; DB 13; Length 717;  
Best Local Similarity 45.4%; Pred. No. 7.5e-101;  
Matches 201; Conservative 78; Mismatches 152; Indels 12; Gaps 6;

QY 5 RNTIHEFKSAKTTLIKIDPALKIKTKVNTADQCANRCTRNRKGLPFTCKAFVDRKARKO 64  
DB 32 RSLANDYQSRKGLVHNEG-GVKQEVQAEIQICAKQCSQD---LLDCRSFDYNNKSQS 86  
QY 65 CLWPFNMSGSGVKKEGHEFDLYENKDYIRNCLIGKGRSYKGTSTKGIKQCPWSSM 124  
DB 87 CRLLPWTQNSNVLLRNQVYDLYQKKDYIRDCVVGNGNTYRTGVTSKTNGRTCQHWRLK 146  
QY 125 IPHSHSLPSSYRGKDLQENYCRNPRGEGGPMWCTSNPEVRVEVCDIPQCSVEECMTCN 184  
DB 147 FPHBKESPTHW--PELEENYCRNPDSDPEGLWCYTTDKNIRHQYCGIKKCEDAVCLTCN 204  
QY 185 GESYRGLMDHTESGICQWRDHQTPHRLKFLPERYPDKGFDNNYCRNPDQCPRPWCYTLD 244  
DB 205 GEDYRGSDRTESGECQWRDLQAPHTHYKPEKYPDKSLDDNYCRNPDSSERPWCYTLD 264  
QY 245 PHTRWECYCAIKTCADNTMNDVDVPLETTECIGOGEGYRGVTNTIANGIPCQRWDSQYPH 304  
DB 265 PNVEREFCHITCKQRISNIEI---TSTCFKEGEGYRGKANTTTSGIPCQRWDIOAPH 321  
QY 305 EHDMTPEFNKCDLRENYCRNPDGSESPWCTTDPNIRVGYCSIPNC-DMSGQDCYRG 363  
DB 322 VHRFLPEKYPCKGLDENYCRNPDGSEAPWCTTTLKNRMAYCFQIKRKTDDVVEPECYHG 381  
QY 364 NGKYNGLNLSQTRSLGTCMMDKNMEDLHRHIFWEPDASKLNENYCRNPDADHGPWCYT 423  
DB 382 NGELYRGVSRKTRGIMCRRWEEKONDLELSLA-OPYLVPLEENYCRNPDSDSHGPWCYT 440  
QY 424 GNPLIPWDCYPTSRCEGDTTPTI 446  
DB 441 MDPNTPFDYCALKPCAGDKVLTIL 463

RESULT 10  
Q902N6 PRELIMINARY; PRT; 709 AA.  
AC Q902N6  
DT 01-DEC-2001 (Tremblrel. 19, Created)

DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)  
DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)  
DE Hepatocyte growth factor-like 1.  
GN HGFLI.  
OS Brachydanio rerio (Zebrafish) (Zebra danio).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;  
OC Cyprinidae; Danio.  
OX NCBI\_TaxID=7955;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Bassett D.I., Wilson S.W.;  
RT "Early expression of zebrafish Hepatocyte Growth Factor-Like 1  
RT suggests a conserved role in vertebrate neural induction.";  
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF370035; AAK54207.1; .  
DR InterPro; IPR000001; Kringle.  
DR InterPro; IPR003014; PAN.  
DR InterPro; IPR001254; Ser.protease\_Try.  
DR Pfam; PF00051; kringle; 4.  
DR Pfam; PF00024; PAN; 1.  
DR Pfam; PF00089; trypsin; 1.  
DR ProDom; PD000395; Kringle; 4.  
DR PROSITE; PS00021; KRINGLE\_1; UNKNOWN\_4.  
DR PROSITE; PS50070; KRINGLE\_2; 4.  
DR PROSITE; PS50240; TRYPsin\_DOM; 1.  
KW Hydrolase; Serine protease.  
SQ SEQUENCE 709 AA; 81271 MW; 9907236C5DB73A20 CRC64;

Query Match 45.0%; Score 1172; DB 13; Length 709;  
Best Local Similarity 46.0%; Pred. No. 1.8e-100;  
Matches 216; Conservative 60; Mismatches 134; Indels 60; Gaps 12;

QY 5 RNTIHEFKSAKTTLIKIDPALKIKTKVNTADQCANRCTRNRKGLPFTCKAF 56  
DB 21 RSLANDYQSRKGLVHNEG-LKSNVNRUTELPRVTEDECAKRCCTSP---ECRAF 69  
QY 57 VFD-KARKQCLWFPF--NSMSGVKKEGHEFDLYENKDYIRNCLIGKGRSYKGTSTK 113  
DB 70 NYEPRSLPCKHLPWVGDNADVRNV--NCDLVEMVYVRKCIYVGKEDYRGKVSTTT 127  
QY 114 SGIKQCPWSSMIPHEHSLPSSYRGKDLQENYCRNPRGEGGPMWCTSNPEVRVEVCDIP 173  
DB 128 SGRTCOQWMSKPEPHDRWTPSATNGLEL--NECRNPDGDRIGPWCYTDPPEVESCNP 185  
QY 174 QCSEVECTMTCGESYRGLMDHTESGICQWRDHQTPHRLKFLPERYPDKGFDNNYCRNPD 233  
DB 186 QCKDEVCITCNGEDYRGQVDHTISGECQWRDQFPQHIIYQPEKYPDKSLDDNYCRNPD 245  
QY 234 GQPRWCYTLDPHTRWECYCAIKTCADNTMNDVDVPLE-----TTECIGOGEGYRGVTN 287  
DB 246 ASPVPWCYTDPMTMERESCDISK-----PEFPARRLRSSYITNCFRAGEDYRGKN 298  
QY 288 TIWNGIPCQRWDSQYPHEDMTPEFNKCDLRENYCRNPDGSESPWCTTDPNIRVGYCS 347  
DB 299 EITSGIPCQRWDAQKPEHPPFKIYEKGLLEENYCRNPDGSEAPWCTTSLPEMRTALCL 358  
QY 348 QIPNC-DMSGQDCYRGKNKNTMGNLSQTRSLGTCMMDKNMEDLHRHIFWEPDASKLN- 405  
DB 359 QIKRCADDIEADCYNEIGKNYRGVVRKTRKILCKQWSIN-----TPHKTKNP 408  
QY 406 -----ENYCRNPDADHGPWCYTGNPLIPWDCYPTSRCEGDTTPTI 446  
DB 409 KTHPEANLTDNYCRNPDGDHGHGWCYTSPKTEFDYCALKQACAGEKVPIT 458

RESULT 11  
Q13208 PRELIMINARY; PRT; 567 AA.  
AC Q13208  
DT 01-NOV-1996 (Tremblrel. 01, Created)  
DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)  
DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)

DE Hepatocyte growth factor-like protein homolog.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20191171; PubMed=10728827;  
 RA Degen S.J.F., McDowell S.A., Waltz S.E., Gould F., Stuart L.A.,  
 RA Carritt B.;  
 RT "Structure of the human DIF15S1A locus: a chromosome 1 locus with 97%  
 RT identity to the chromosome 3 gene coding for hepatocyte growth factor-  
 RT like protein.";  
 RL DNA Seq. 8:409-413(1998).  
 DR EMBL; U28054; AAC63092.1; -.  
 DR HSP; P00747; 2PK4.  
 DR MEROPS; S01.977; -.  
 DR InterPro; IPR000001; Kringle.  
 DR InterPro; IPR003014; PAN.  
 DR InterPro; IPR003609; Pan\_app.  
 DR Pfam; PF00051; kringle; 4.  
 DR Pfam; PF00024; PAN; 1.  
 DR Pfam; PF00089; trypsin; 1.  
 DR PRINTS; PR00018; KRINGLE.  
 DR PRODOM; PD000395; Kringle; 4.  
 DR SMART; SM00130; KR; 4.  
 DR SMART; SM00473; PAN\_AP; 1.  
 DR PROSITE; PS00021; KRINGLE\_1; 2.  
 DR PROSITE; PS00070; KRINGLE\_2; 4.  
 DR PROSITE; PS02040; TRYPSIN\_DOM; 1.  
 DR Hydrolase; Serine protease.  
 KW Hydrolyase; Serine protease.  
 SQ SEQUENCE 567 AA; 64117 MW; 3FC38B07F1645810 CRC64;  
 Query Match 44.78; Score 1163.5; DB 4; Length 567;  
 Best Local Similarity 47.18; Pred. No. 8.2e-100;  
 Matches 196; Conservative 62; Mismatches 149; Indels 9; Gaps 4;  
 Qy 33 VNTADQCANRTRKGLPFTCKAFVDFKARKQCLWFPNSMSGVKKFEGHEFDLYENKD 92  
 Db 25 VADAEACAGRC---GLLMDCAWFHYNVSSHGCGQLLPWQHSPHSLRHSRCGLDFQKDD 80  
 Qy 93 YIRNCITGKSGYKGTVSITKSGIKQCPWSSMIPHEHSFLPSSYRGKDLQENYCRNPRGE 152  
 Db 81 YIRTCIMNNGVGYRDTMTATVGLSCQAWSHKFPNDHQYMTPLRNG--LEENFCRNPDGD 138  
 Qy 153 EGGPWCFSTNSPEVRYEYCDIPQCSVEVCMTCNGESYRGMLDHTESGKTCQWDHQTPIRH 212  
 Db 139 PGGPWCHTDDPAVRFSQGIKSLVAACVWCNGEYRGAVDRTESGRCQWDLQHPHQH 198  
 Qy 213 KFLPERYPDKGDFDNYCRNPDGQPRPWCYTLDPHTRWEYCAIKTCADNTMDTIDVPLETT 272  
 Db 199 PFEKGKFLDQGLDDNYCRSPDGSQRPWCYTDPQIEREFCDLPRGSEAPQREA--TSV 256  
 Qy 273 ECIQGGEGYRGTVNTWINGIPCORWDSQYPEHDMTPENFKCKDLRENYCRNPDGSGSP 332  
 Db 257 SCFRGKGEGYRGTVNTTAGVPCQWDAQIPQHRFTPEKYACKDLRENFRCNPDGSEAP 316  
 Qy 333 WCFTDPNIRVYCSQIPNC--DMSGHGDQCYRGNGKNYGNLSQTSGLTCSMDKKNMEDL 391  
 Db 317 WCFTLPRTGRVFCYQIRRCRDDVVRPDQCYHGAGEYRGTVSKTRKGVCQCRWSAETPHK 376  
 Qy 392 HRIHFWEPDASKLNEYCRNPDGDAHPWCYTGNPLIPWDYCPISRCBGDTPTTIV 447  
 Db 377 PQFTFTSEPHAQLEENFCQDGDGSHGWCYTMDPRTPFDYCALRCADDQPPSIL 432  
 RESULT 12  
 ID P70521 PRELIMINARY; PRT; 716 AA.  
 AC P70521;  
 DT 01-FEB-1997 (TrEMBLrel. 02, Created)  
 DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)

DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)  
 DE Macrophage stimulating protein precursor.  
 GN MSP.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=97011126; PubMed=8858136;  
 RA Onshiro K., Iwama A., Matsuno K., Ezaki T., Sakamoto O., Hamaguchi I.,  
 RA Takasu N., Suda T.;  
 RT "Molecular cloning of Rat Macrophage-stimulating protein and its  
 RT involvement in the Male Reproductive System.";  
 RL Biochem. Biophys. Res. Commun. 227:273-280(1996).  
 CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE  
 CC TRYPSIN FAMILY.  
 DR EMBL; X95096; CAA64473.1; -.  
 DR HSP; P00747; IKRN.  
 DR MEROPS; S01.975; -.  
 DR InterPro; IPR001314; Chymotrypsin.  
 DR InterPro; IPR000001; Kringle.  
 DR InterPro; IPR003014; PAN.  
 DR InterPro; IPR003609; Pan\_app.  
 DR Pfam; PF00051; kringle; 4.  
 DR Pfam; PF00024; PAN; 1.  
 DR Pfam; PF00089; trypsin; 1.  
 DR PRINTS; PR00722; CHYMOTRYPSIN.  
 DR PRINTS; PR00018; KRINGLE.  
 DR PRODOM; PD000395; Kringle; 4.  
 DR SMART; SM00130; KR; 4.  
 DR SMART; SM00473; PAN\_AP; 1.  
 DR SMART; SM00020; Tryp\_Spc; 1.  
 DR PROSITE; PS00021; KRINGLE\_1; 4.  
 DR PROSITE; PS00070; KRINGLE\_2; 4.  
 DR PROSITE; PS02040; TRYPSIN\_DOM; 1.  
 DR Hydrolase; Serine protease; Signal.  
 KW SIGNAL  
 FT SIGNAL 1 31 POTENTIAL.  
 SQ SEQUENCE 716 AA; 80733 MW; 06B7DF3EF56D921F CRC64;  
 Query Match 43.9%; Score 1144; DB 11; Length 716;  
 Best Local Similarity 46.18; Pred. No. 7.2e-98;  
 Matches 196; Conservative 70; Mismatches 141; Indels 18; Gaps 6;  
 Qy 33 VNTADQCANRTRKGLPFTCKAFVDFKARKQCLWFPNSMSGVKKFEGHEFDLYENKD 92  
 Db 50 VADAEACAGRC---GPLLDCRAFHYNMSSHGCGQLLPWQHSLRAQLHSSLCGLDFQKDD 105  
 Qy 93 YIRNCITGKSGYKGTVSITKSGIKQCPWSSMIPHEHSFLPSSYRGKDLQENYCRNPRGE 152  
 Db 106 YVTCIMDNGASRGTVARTADGLPCQAWSRFRPNHDKYTPPKNG--LEENFCRNPDGD 163  
 Qy 153 EGGPWCFSTNSPEVRYEYCDIPQCSVEVCMTCNGESYRGMLDHTESGKTCQWDHQTPIRH 212  
 Db 164 PRGPWCYTNSVRFSQGIKSCREAVCWCNGEDYRGEDVDTESGRCQWDLQHPHSH 223  
 Qy 213 KFLPERYPDKGDFDNYCRNPDGQPRPWCYTLDPHTRWEYCAIKTCADN-----TMDT 265  
 Db 224 PFHPEKFPDKALKONCYCRNPDASERPWCYTDDPNVEREFCDLPCGPNLPPTTKGSKSQ 283  
 Qy 266 DVPLETTEICQGGEGYRGTVNTWINGIPCORWDSQYPEHDMTPENFKCKDLRENYCRN 325  
 Db 284 RNKYKASNCFRGKGEDYRGTVNTTSAGVPCQWDAQNPHQHRFVPEKYACKDLRENFERN 343  
 Qy 326 PDGSESPWCFTTDPNIRVYCSQIPNC--DMSGHGDQCYRGNGKNYGNLSQTSRGLTCSMW 384  
 Db 344 PDGSEAPWCFTSRPGLRVAFYCIQIPRCTEEVVPVPECYHGSGEQYRGVSKTRKGVCQHW 403  
 Qy 385 DKNMEDLHRIHFWEPD--SKLNEYCRNPDGDAHPWCYTGNPLIPWDYCPISRCBGDT 442  
 Db 404 --SSETPHKPQFTPTSAPHAGLEANFCRNPDGDSHGWCYTLDPELTLDYCALKRCDDQ 461



QY 149 PRGEGGWCFTSNPEVRYEVCDIPQCEVEECMTNGESYRGLMDHTESGKTCQWRDHT 208  
 Db 155 PNDQGGWCFTTDPQRYEYCNIECEB-ECMYSGEKIEGKISKTMSGLDCQWDSOS 213  
 QY 209 PHRHFELPERYDKGFDNDYCNPNQDQRPWCYTLDPHTRWEYCAIKTCADTMNDTVP 268  
 Db 214 PHAGYIIPAKFPSKMLKNYCNPNQDQRPWCFTTDPNKRWEYCDIPRCT-----TPPPPP 269  
 QY 269 LETTECICOGGEGYRGTVNTIWNIGIPQORWDSQYRPHEDMTENFKCKDLNRYCNPNPDG 328  
 Db 270 GFTYOCCKGRGENYRGTVSVTASGKTCQWSEQTPHRHRTPEPFCKNLEENYCNPNPDG 329  
 QY 329 SESPCFTTDPNIRVGYCSQIPNC-----DMSHG-----ODCYRGNGKNYMGNL 372  
 Db 330 ETAPWCYTTDSLRWEYC-EIPSCSSVSPDQSSVLPQETPVVQECYQCGNGKSYRGTS 388  
 QY 373 SOTRSGLCYSMDKMDLHRHI---FWEPDASKLENYCNPNPDQDAGPWCYTCNPLIP 429  
 Db 389 SYTNTGKQCSWSTPHSHSKTANF--PDAG-LEMNYCNPNPDQDQRPWCFTTDPSPVR 445  
 QY 430 WDCYPISRC 438  
 Db 446 WEYCNLKR 454

RESULT 15

O18783  
 ID O18783 PRELIMINARY; PRT; 806 AA.  
 AC O18783;  
 DT 01-JAN-1998 (TRENBLrel. 05, Created)  
 DT 01-JAN-1998 (TRENBLrel. 05, Last sequence update)  
 DT 01-MAR-2002 (TRENBLrel. 20, Last annotation update)  
 DE Plasminogen.  
 OS Macropus eugenii (Tamar wallaby).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Metatheria; Diprotodontia; Macropodidae; Macropus.  
 OX NCBI\_TaxID=9315;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=LIVER;  
 RX MEDLINE=98004511; PubMed=9342350;  
 RA Lawt R.M., Schwartz K., Patthy L.;  
 RT "Convergent evolution of apolipoprotein(a) in primates and hedgehog."  
 RL Proc. Natl. Acad. Sci. U.S.A. 94:11992-11997(1997).  
 CG -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE  
 CC TRYPSIN FAMILY.  
 DR EMBL; AF012297; AAB65760.1; -;  
 DR HSP; P00747; 5HPG.  
 DR MEROPS; S01.233;  
 DR InterPro; IPR001314; Chymotrypsin.  
 DR InterPro; IPR000001; Kringle.  
 DR InterPro; IPR003014; PAN.  
 DR InterPro; IPR003609; Pan\_app.  
 DR InterPro; IPR001254; Ser\_protease\_Try.  
 DR Pfam; PF00051; kringle; 5.  
 DR Pfam; PF00024; PAN; 1.  
 DR Pfam; PF00089; trypsin; 1.  
 DR PRINTS; PR00722; CHYMOTRYPSIN.  
 DR PRINTS; PR00018; KRINGLE.  
 DR PRODOM; PD000395; Kringle; 5.  
 DR SMART; SM00130; KR; 4.  
 DR SMART; SM00473; PAN\_AP; 1.  
 DR SMART; SM00020; TRY-SPC; 1.  
 DR PROSITE; PS00021; KRINGLE.1; 5.  
 DR PROSITE; PS00070; KRINGLE.2; 5.  
 DR PROSITE; PS0240; TRYPSIN\_DOM; 1.  
 DR PROSITE; PS00134; TRYPSIN\_SER; 1.  
 DR PROSITE; PS00135; TRYPSIN\_SER; 1.  
 KW Hydrolase; Serine protease.  
 SQ SEQUENCE 806 AA; 90981 MW; 95FAA86DC20064D5 CRC64;

Query Match 41.5%; Score 1080.5; DB 6; Length 806;  
 Best Local Similarity 42.8%; Pred. No. 6.8e-92;

Matches 191; Conservative 79; Mismatches 143; Indels 33; Gaps 11;  
 QY 20 IKIDPALKTKTKK-----VNTADQCANRCTRNKGLPETCKAFVDFDKARKOCLAFPPFNMS 75  
 Db 26 IKTEGASLSNOKKQFVASSTEECALC--EKETEFCRSFEHYNKKEQKCVIMSENKSTS 83  
 QY 76 GYKKEFGHEFDLYENKDYIRNCIIGKGRSYKGTVSITKSGIKCQWPSSMIPHEHSLPSS 135  
 Db 84 SVERRK--RDVLFKEKRIYLSLDCSKSGNGRNYRGLTSLKTSKITCQKWSLSPHPVNPYAPSK 141  
 QY 136 YGKDLQENYCNPNRGEEGGPGWCFTSNPEVRYEVCDIPQCEVEECMTNGESYRGLMDHT 195  
 Db 142 YPDAGLEKNYCNPNDDDVKGPWCYTTNPDRIYECVPEEC-EDCEMHCSENGYRGITISKT 200  
 QY 196 ESGKICQWRDHTPHRHKFLPERYPDKGFDNDYCNPNPDGPRPWCYTLDPHTRWEYCAIK 255  
 Db 201 ESGIECPWDSQEPHSHEYIPSKFPSKDLKENYCNPNPDGEPRPWCFTSNPEKWEFCNIP 260  
 QY 256 TCADNTMNDTVDPLETTECICOGGEGYRGTVNTIWNIGIPQORWDSQYRPHEDMTENFKC 315  
 Db 261 RCS-----SPPPPPGPMQLCKLGRGENYRGKIAVTKSGHTCQRWNKQTPHKHNRTPENFPC 316  
 QY 316 KDLRENYCNPNPDGSESPWCFTTDPNIRVGYCSQIPNCDSHG-----ODCYRGN 364  
 Db 317 RGLDENYCNPNPDGELEPWCYTTNPDVROEYCA-IPSCGTSSPHTDRVEQSPVIOECYEKG 375  
 QY 365 GKNYMGNLQSOTRSGLTCSMDKNMEDLHRHIFWEPD---ASKLNEYCNPNPDQDAGHPWC 421  
 Db 376 GENYRGTTSTTISGKKCQAWSSMTPHQHK---TPDNFFNADLIRNYCNPNPDG-KSPWC 431  
 QY 422 YTGPNLIPWDYCPISRCGDTTPTIV 447  
 Db 432 YTMDDPTVRWEFCNLEKCSG-TGSTVL 456

Search completed: April 25, 2003, 15:51:43  
 Job time : 40 secs

**THIS PAGE BLANK (USPTO)**